

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2001, 13:22:11 ; Search time 18.61 Seconds  
(without alignments)  
1778.740 Million cell updates/sec

Title: US-09-689-430-2

Perfect score: 7869

Sequence: 1 M0E1STCFLLCFLRRCFSA.....WVHQLRMVGLCEAODLY (1471)

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

1: Issued\_Patents\_AA:\*  
2: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep:\*  
6: /cgn2\_6/ptodata/2/1aa/6C.COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7857	99.8	1471	1 US-08-683-839B-3	Sequence 3, Appl1
2	7616	96.8	1661	2 US-08-882-083-2	Sequence 2, Appl1
3	7616	96.8	1661	2 US-08-558-107-2	Sequence 2, Appl1
4	7616	96.8	1661	4 US-09-243-539-2	Sequence 2, Appl1
5	7419	94.3	2351	1 US-08-121-202-2	Sequence 2, Appl1
6	7410	94.2	2351	6 5422260-1	Patent No. 5422260
7	7406	94.1	2351	6 5171844-2	Patent No. 5171844
8	7403	94.1	2351	6 5171844-2	Patent No. 5171844
9	7315	93.0	2332	1 US-07-864-004B-4	Sequence 4, Appl1
10	7315	93.0	2332	1 US-08-251-937A-2	Sequence 4, Appl1
11	7315	93.0	2332	1 US-08-212-133A-2	Sequence 2, Appl1
12	7315	93.0	2332	1 US-08-474-503-2	Sequence 2, Appl1
13	7315	93.0	2332	2 US-08-670-707A-2	Sequence 2, Appl1
14	7315	93.0	2332	4 US-09-037-601-2	Sequence 2, Appl1
15	7315	93.0	2332	5 PCT-US93-03275-4	Sequence 4, Appl1
16	7315	93.0	2332	5 PCT-US94-13200-2	Sequence 4, Appl1
17	7296	92.7	2332	1 US-08-276-594A-2	Sequence 2, Appl1
18	7276	92.5	2332	4 US-09-324-867-3	Sequence 3, Appl1
19	6552	83.3	1443	2 US-08-670-707A-39	Sequence 3, Appl1
20	6552	83.3	1443	4 US-09-037-601-39	Sequence 39, Appl1
21	6401	81.3	2343	4 US-09-324-867-2	Sequence 39, Appl1
22	6347	80.7	2133	2 US-08-670-707A-37	Sequence 37, Appl1
23	6347	80.7	2133	4 US-09-037-601-37	Sequence 37, Appl1
24	6301	80.1	2319	1 US-08-212-133A-8	Sequence 8, Appl1
25	6301	80.1	2319	1 US-08-474-503-6	Sequence 6, Appl1
26	6301	80.1	2319	2 US-08-670-707A-6	Sequence 6, Appl1
27	6301	80.1	2319	4 US-09-037-601-6	Sequence 6, Appl1

28	6301	80.1	2319	5 PCT-US94-13200-6	Sequence 6, Appl1
29	6237.5	79.3	2115	4 US-09-324-867-5	Sequence 5, Appl1
30	6196	78.7	2304	4 US-09-324-867-4	Sequence 4, Appl1
31	2444.5	31.1	341	1 US-08-121-202-4	Sequence 4, Appl1
32	2394.5	30.4	2183	3 US-08-746-111-5	Sequence 5, Appl1
33	1726	21.9	368	1 US-08-212-133A-6	Sequence 6, Appl1
34	1726	21.9	368	1 US-08-474-503-4	Sequence 4, Appl1
35	1726	21.9	368	2 US-08-670-707A-4	Sequence 4, Appl1
36	1726	21.9	368	4 US-09-037-601-4	Sequence 4, Appl1
37	1726	21.9	368	5 PCT-US94-13200-4	Sequence 4, Appl1
38	1722	21.9	367	1 US-07-864-004B-2	Sequence 2, Appl1
39	1722	21.9	367	1 US-08-251-937A-2	Sequence 2, Appl1
40	1722	21.9	367	5 PCT-US93-03275-2	Sequence 2, Appl1
41	1153	14.7	218	1 US-07-607-538C-5	Sequence 5, Appl1
42	1153	14.7	218	2 US-08-162-402B-5	Sequence 5, Appl1
43	846	10.8	160	2 US-08-162-402B-15	Sequence 15, Appl1
44	824	10.5	154	2 US-08-162-402B-17	Sequence 17, Appl1
45	650	8.3	463	2 US-08-162-402B-9	Sequence 9, Appl1

#### ALIGNMENTS

```
RESULT 1
US-08-683-839B-3
; Sequence, 3, Application US/08683839B
; Patent No. 5744326
; GENERAL INFORMATION:
; APPLICANT: ILL, Charles . R. et al.
; TITLE OF INVENTION: Use of Viral Cis-Acting Post-Transcriptional
; TITLE OF INVENTION: Regulatory Sequences To Increase Expression of
; TITLE OF INVENTION: Intronsless Genes Containing Near-Consensus Splice Sites
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/683,839B
; FILING DATE: 11-MARCH-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Remillard, Jane E.
; REGISTRATION NUMBER: 38,872
; REFERENCE/DOCKET NUMBER: TTI-138
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1471 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-683-839B-3

Query Match 99.8%; Score 7857; DB 1; Length 1471;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1469; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 1 M0E1STCFLLCFLRRCFSAATRRYIGAVELSDMYQSDLGELPVDARFPVPRPKSPFN 60
```







Matches 1450; Conservative 5; Mismatches 10; Indels 202; Gaps 4;

```

OY 1 MOELSTCFELCLRECFATRRYYLGAVELSMDYMOGDLGELPYDARPPVPKSPFN 60
Db 1 MEELSTCFELCLRECFATRRYYLGAVELSMDYMOGDLGELPYDARPPVPKSPFN 60
OY 61 TSVVYKKTLEVEFTVHLFNIAKPRPPMGLGPTIOAEVYDVTITLKNMASHVPSLHAY 120
Db 61 TSVVYKKTLEVEFTVHLFNIAKPRPPMGLGPTIOAEVYDVTITLKNMASHVPSLHAY 120
OY 121 GVSYKASGAGYDQTSOREKEDKVPFGSGSHYVVOYLKNGPMASDPLCLTYSLSH 180
Db 121 GVSYKASGAGYDQTSOREKEDKVPFGSGSHYVVOYLKNGPMASDPLCLTYSLSH 180
OY 181 VDLVMDNSGLIGALLVREGSLAKEKTOYLHKFTLLFAVPDEGKSMSEFNKSLMORD 240
Db 181 VDLVMDNSGLIGALLVREGSLAKEKTOYLHKFTLLFAVPDEGKSMSEFNKSLMORD 240
OY 241 AASARAMPKMTVNGVYVRSJLPLGICHRKSYVNHVIGMTTPEVHSITFLEGHTEFLVNH 300
Db 241 AASARAMPKMTVNGVYVRSJLPLGICHRKSYVNHVIGMTTPEVHSITFLEGHTEFLVNH 300
OY 301 RASLEISPIITFLTAOTLLMDLGOFLCCHSHOHODGMEAVKYVDSCEPPEOLMKNE 360
Db 301 RASLEISPIITFLTAOTLLMDLGOFLCCHSHOHODGMEAVKYVDSCEPPEOLMKNE 360
OY 361 EAEYDDDLTDESEMDVFEEDDONSFIQISVAKKHPTWVHYTAAEEBMDVAPLVA 420
Db 361 EAEYDDDLTDESEMDVFEEDDONSFIQISVAKKHPTWVHYTAAEEBMDVAPLVA 420
OY 421 PDDRSYKQOYLNNGBORIGRKRYKRVFMAVDETFKTRREALQIHESGILGPLLYGEVDTL 480
Db 421 PDDRSYKQOYLNNGBORIGRKRYKRVFMAVDETFKTRREALQIHESGILGPLLYGEVDTL 480
OY 481 LIIFENQASRPNTYPHGITTDRPLYSRRLKGVKXHLKDFPLPEGEIRKYVTVVEDGP 540
Db 481 LIIFENQASRPNTYPHGITTDRPLYSRRLKGVKXHLKDFPLPEGEIRKYVTVVEDGP 540
OY 541 TKSDFRCLTRYSSFVNNERDLASGLIPLLICKEYSDOGNOIMSKRWVILFESVDE 600
Db 541 TKSDFRCLTRYSSFVNNERDLASGLIPLLICKEYSDOGNOIMSKRWVILFESVDE 600
OY 601 NRSWLTENIORFLPNPAGVLEDEPEQASNIMHSINGVPSIQLSYCLHEVAWYILS 660
Db 601 NRSWLTENIORFLPNPAGVLEDEPEQASNIMHSINGVPSIQLSYCLHEVAWYILS 660
OY 661 IGAOTDFLSVFSSGYTFHAKVYEDTLTLPFPGSETVMSMENPGMLTILGCHNSDFRNRG 720
Db 661 IGAOTDFLSVFSSGYTFHAKVYEDTLTLPFPGSETVMSMENPGMLTILGCHNSDFRNRG 720
OY 721 MTALLKAVSSC---DKNTGDDY-----EDSYEDISAVLSKNNATPRFSQNSRRPS 769
Db 721 MTALLKAVSSC---DKNTGDDY-----EDSYEDISAVLSKNNATPRFSQNSRRPS 769
OY 770 TROKOFNAT----- 778
Db 770 TROKOFNAT----- 778
OY 775 TROKOFNATLIPENDIEKTDPMFAHRTPMKIQNVSSSDDLMLLRQSTPHGLSLSDLOE 834
Db 775 TROKOFNATLIPENDIEKTDPMFAHRTPMKIQNVSSSDDLMLLRQSTPHGLSLSDLOE 834
OY 779 ----- 778
Db 779 ----- 778
OY 835 AKYTFSDPSPGALDSNNSLSEMTHFRPOLHHSQDMVFPESGLOJRLNEKLGTTADPL 894
Db 835 AKYTFSDPSPGALDSNNSLSEMTHFRPOLHHSQDMVFPESGLOJRLNEKLGTTADPL 894
OY 895 AMDNHGTQIKKEEMKSOEKSPKTAFFKKDPTIISLACESNHAIAAINESQNKPELEVT 954
Db 895 AMDNHGTQIKKEEMKSOEKSPKTAFFKKDPTIISLACESNHAIAAINESQNKPELEVT 954
OY 779 -----PVLKRHQREITRTTLOSQDEIDYDPTISVENKKEDFDYDEDEN 824
Db 779 -----PVLKRHQREITRTTLOSQDEIDYDPTISVENKKEDFDYDEDEN 824
OY 955 WAKOGRTERLCSQNPVYLKRHQREITRTTLOSQDEIDYDPTISVENKKEDFDYDEDEN 1014
Db 955 WAKOGRTERLCSQNPVYLKRHQREITRTTLOSQDEIDYDPTISVENKKEDFDYDEDEN 1014
OY 825 OSPRSFOKTRHYFLAVERLMDYGMSSSPVYLNRRASGSGVPQFKVYVFOEFTDGSFTQ 884
Db 825 OSPRSFOKTRHYFLAVERLMDYGMSSSPVYLNRRASGSGVPQFKVYVFOEFTDGSFTQ 884
OY 1015 OSPRSFOKTRHYFLAVERLMDYGMSSSPVYLNRRASGSGVPQFKVYVFOEFTDGSFTQ 1074
Db 1015 OSPRSFOKTRHYFLAVERLMDYGMSSSPVYLNRRASGSGVPQFKVYVFOEFTDGSFTQ 1074

```

```

OY 885 PLYRGELNEHGLIGPYTIRAEVEDINIVTFPRNOASRPVSPVSSLTISEDDOROGAEPKRN 944
Db 1075 PLYRGELNEHGLIGPYTIRAEVEDINIVTFPRNOASRPVSPVSSLTISEDDOROGAEPKRN 1134
OY 945 FVKPNETKYTFWKVVOHNAAPTKEDEFCAMAYFSDVDLEKDVHSGSLGPLLVCHTNTLNP 1004
Db 1135 FVKPNETKYTFWKVVOHNAAPTKEDEFCAMAYFSDVDLEKDVHSGSLGPLLVCHTNTLNP 1194
OY 1005 AHGROYVQOERALFTTIDETKSWTFTENMERNCRAPCNIOEMEDPTFKENYRFAINCYI 1064
Db 1195 AHGROYVQOERALFTTIDETKSWTFTENMERNCRAPCNIOEMEDPTFKENYRFAINCYI 1254
OY 1065 MDTLPLGLVMAQDORIRWYLLSNGSMENHSHIFSGHVTFVKKREYKNAALNLYPGVET 1124
Db 1255 MDTLPLGLVMAQDORIRWYLLSNGSMENHSHIFSGHVTFVKKREYKNAALNLYPGVET 1314
OY 1125 VEMLPKAGIRVRECLIGENHLAGKSTLFLVYSNKCQPLGMAAGHIDFOITASGOYGO 1184
Db 1315 VEMLPKAGIRVRECLIGENHLAGKSTLFLVYSNKCQPLGMAAGHIDFOITASGOYGO 1374
OY 1185 WAPKLARLHYSGSINAMSTKEPFSWIKVDLAPMIIHGIKTQGAROKFSSLYISOITMY 1244
Db 1375 WAPKLARLHYSGSINAMSTKEPFSWIKVDLAPMIIHGIKTQGAROKFSSLYISOITMY 1434
OY 1245 SLDDKKMOTYRGNSGTLMVFEFGVNDSSGIKHNIFNPITIRYIRLHPHTHSISTLME 1304
Db 1435 SLDDKKMOTYRGNSGTLMVFEFGVNDSSGIKHNIFNPITIRYIRLHPHTHSISTLME 1494
OY 1305 LMGCDLNSCMLPGLMESKAISDAQITASSYFTNMFATSPSKARLHLOGRSNANRPVNN 1364
Db 1495 LMGCDLNSCMLPGLMESKAISDAQITASSYFTNMFATSPSKARLHLOGRSNANRPVNN 1554
OY 1365 PKEMLOVDFOKTMKYVTGTOGVKSLTSMYKVEFLISSODGHOVTLFONGKVKVYFOG 1424
Db 1555 PKEMLOVDFOKTMKYVTGTOGVKSLTSMYKVEFLISSODGHOVTLFONGKVKVYFOG 1614
OY 1425 NODSFTPVVNSLDPLRLRYLRHPHOSVHOIALRMEVLGEADLY 1471
Db 1615 NODSFTPVVNSLDPLRLRYLRHPHOSVHOIALRMEVLGEADLY 1661

--
US-058-107-2
: Result 3
: Sequence 2, Application US/08558107
: Patent No. 5910481
: GENERAL INFORMATION:
: APPLICANT: VOORBERG, Johannes J.
: TITLE OF INVENTION: HYBRID PROTEINS WITH MODIFIED ACTIVITY
: NUMBER OF SEQUENCES: 17
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Foley & Lardner
: STREET: 3000 K Street, N.W., Suite 500
: CITY: Washington
: STATE: D.C.
: COUNTRY: USA
: ZIP: 20007-5109
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/558,107
: FILING DATE: 13-NOV-1995
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: ISACSON, John P.
: REGISTRATION NUMBER: 33,715
: REFERENCE/DOCKET NUMBER: 30472/212
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (202)672-5300
: TELEFAX: (202)672-5399

```



```

; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1661 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-558-107-2

```

```

Query Match      96.8%: Score 7616; DB 2; Length 1661;
Best Local Similarity 87.0%: Fred. No. 0;
Matches 1450; Conservative 5; Mismatches 10; Indels 202; Gaps 4;

```

```

QY 1 MEIETSTCFCLLRFCSATRRYLGAVELSMQSDGLGELPVDPARPPRPVPSFPPN 60
D 1 MEIETSTCFCLLRFCSATRRYLGAVELSMQSDGLGELPVDPARPPRPVPSFPPN 60
QY 61 TSVYKKTLEVEFTVHLFNIAKPRPPMGLGPTIOAEVYDTVTITLKNMASHPVSLAHV 120
D 61 TSVYKKTLEVEFTVHLFNIAKPRPPMGLGPTIOAEVYDTVTITLKNMASHPVSLAHV 120
QY 121 GVSYWKASGEAEYDQTSOREKEDKVPFGSGHTYVMOYLKENGPMASPLCLTYSLSH 180
D 121 GVSYWKASGEAEYDQTSOREKEDKVPFGSGHTYVMOYLKENGPMASPLCLTYSLSH 180
QY 181 VDLVKDLNSGLIGALLVCREGLAKEKTOTLHKFILLFAVPEBGKSWHSETKNSLMQDD 240
D 181 VDLVKDLNSGLIGALLVCREGLAKEKTOTLHKFILLFAVPEBGKSWHSETKNSLMQDD 240
QY 241 AASARAPKMTVNGVYVNRSLPGLIGCHRSYVWHVIGMTTPEVHSIFLEGHTEFLVRNH 300
D 241 AASARAPKMTVNGVYVNRSLPGLIGCHRSYVWHVIGMTTPEVHSIFLEGHTEFLVRNH 300
QY 301 RQASLEIPITFLTAOTLMDLGOFLFCGISHQDGEAYKVDSCBEPQLRKNNK 360
D 301 RQASLEIPITFLTAOTLMDLGOFLFCGISHQDGEAYKVDSCBEPQLRKNNK 360
QY 361 EAEYDQDLTSEMVDVFRDDNSPSFIOIRSVAKKHPTWYHIAEEDMDYAPLVLA 420
D 361 EAEYDQDLTSEMVDVFRDDNSPSFIOIRSVAKKHPTWYHIAEEDMDYAPLVLA 420
QY 421 PDDRSYKQYLYNNGPQIRGRKYYKVFMAATDEFTKREAIQIHESGILGPLLYGEVDTL 480
D 421 PDDRSYKQYLYNNGPQIRGRKYYKVFMAATDEFTKREAIQIHESGILGPLLYGEVDTL 480
QY 481 LIIRKQASRPVNIYRPHGTDVRLYXSRPLPKGVNHLKDFPLPGELFKYKMTVVEDGP 540
D 481 LIIRKQASRPVNIYRPHGTDVRLYXSRPLPKGVNHLKDFPLPGELFKYKMTVVEDGP 540
QY 541 TKSDPRCLTRYYSFVNNERDLASGLIGPLLCYKESVDQRNQIMSDKRNILFESVPE 600
D 541 TKSDPRCLTRYYSFVNNERDLASGLIGPLLCYKESVDQRNQIMSDKRNILFESVPE 600
QY 601 NRSWYLTENIORFLPNPAGVQLEDEPQASNIHMSINGVDSLOISVCLHEVAIYIIS 660
D 601 NRSWYLTENIORFLPNPAGVQLEDEPQASNIHMSINGVDSLOISVCLHEVAIYIIS 660
QY 661 IGAQTDLFISVFESGTYFKHKMYEDTLTFPSSGETVFMSMNPGLMILGCHNSDRNRG 720
D 661 IGAQTDLFISVFESGTYFKHKMYEDTLTFPSSGETVFMSMNPGLMILGCHNSDRNRG 720
QY 721 MTALLKAVSSCIPREGEDDYDLLEKIFSEDYDI-----VDSILFERFSQNSRHPS 774
D 721 MTALLKAVSSCIPREGEDDYDLLEKIFSEDYDI-----VDSILFERFSQNSRHPS 774
QY 770 TRQKOFNAT----- 778
D 770 TRQKOFNAT----- 778
QY 775 TRQKOFNATITPENDIEKTDPMFAHRTMPKIQNVSSDMLMLRQSPPHGLSLDLQ 834
D 775 TRQKOFNATITPENDIEKTDPMFAHRTMPKIQNVSSDMLMLRQSPPHGLSLDLQ 834
QY 779 ----- 778
D 779 ----- 778
QY 835 AKYETFSDDPSGALIDSNNSISEMTHFRPOLHHSQDMVFTPESGLQLRLNEKLGTTADPL 894
D 835 AKYETFSDDPSGALIDSNNSISEMTHFRPOLHHSQDMVFTPESGLQLRLNEKLGTTADPL 894

```

```

QY 779 ----- 778
D 895 AMDNHVGTQIPKEEMKSOEKSPKTAFAKKKDTILSLNACESNHAIAINEGONKPEIEVT 954
QY 779 ----- 778
D 955 WAKOGTERLCSQNPVPLKRHOEITRTTLOSQDEEIDYDQTSYEMKEDFDIYDEDEN 1014
QY 825 OSPRSFQKTRHYFTAAYERLMDYGMSSSPHYLRRAQSGSVPOKKAFFOEFIDGSEFTO 884
D 1015 OSPRSFQKTRHYFTAAYERLMDYGMSSSPHYLRRAQSGSVPOKKAFFOEFIDGSEFTO 1074
QY 885 PLYRGELENLGLPYTIAEVEDNIMYTFRNQASRPYSFYSLSIYEDDROGAEPKRN 944
D 1075 PLYRGELENLGLPYTIAEVEDNIMYTFRNQASRPYSFYSLSIYEDDROGAEPKRN 1134
QY 945 FVKPNETKTYFVKVQHHNAPTDQEDCKAMAYFSVDLEKDYHSLIGPLLVCHTNLNP 1004
D 1135 FVKPNETKTYFVKVQHHNAPTDQEDCKAMAYFSVDLEKDYHSLIGPLLVCHTNLNP 1194
QY 1005 AHGRQYVOEFALFTIPDETQSWYFTENMERNCRAPCNIQMEDPTKEKNTYRPHANGYI 1064
D 1195 AHGRQYVOEFALFTIPDETQSWYFTENMERNCRAPCNIQMEDPTKEKNTYRPHANGYI 1254
QY 1065 MDTLPLVMAODORTIRWYLLSMGSENENIHSIFSGHVFYVRKKEEYKALNYLPGVEET 1124
D 1255 MDTLPLVMAODORTIRWYLLSMGSENENIHSIFSGHVFYVRKKEEYKALNYLPGVEET 1314
QY 1125 VEMLPKSGATWRECELGELHAGMSTLYFLVSNKQCPPLCMASGHIRDPOITYASQYQ 1184
D 1315 VEMLPKSGATWRECELGELHAGMSTLYFLVSNKQCPPLCMASGHIRDPOITYASQYQ 1374
QY 1185 WAPKLARLHYSGSINAMSKPEFSMLKYDLPMTIHIKIQAGROKSSLYISQITMY 1244
D 1375 WAPKLARLHYSGSINAMSKPEFSMLKYDLPMTIHIKIQAGROKSSLYISQITMY 1434
QY 1245 SLDGKQWQYRNGNSTGTLLWVFGNVDSSGIKHINFPILARYIRLHPTHYRSINSTLME 1304
D 1435 SLDGKQWQYRNGNSTGTLLWVFGNVDSSGIKHINFPILARYIRLHPTHYRSINSTLME 1494
QY 1305 LMGCDLNSCMLPGMESKASIDAOITASSYFTNMATSPSKARLHLQGRNAMPPOVNN 1364
D 1495 LMGCDLNSCMLPGMESKASIDAOITASSYFTNMATSPSKARLHLQGRNAMPPOVNN 1554
QY 1365 PKEMLOVDFOKTMKYTGVTGCVKSLTSMYKEFLISSODGHOWTLFPONGKVKVFOG 1424
D 1555 PKEMLOVDFOKTMKYTGVTGCVKSLTSMYKEFLISSODGHOWTLFPONGKVKVFOG 1614
QY 1425 NQDSFTPVVNSLDPPILTRYLRIRHPOSVWQIALRMEVYLGCEAODLY 1471
D 1615 NQDSFTPVVNSLDPPILTRYLRIRHPOSVWQIALRMEVYLGCEAODLY 1661

RESULT 4
US-09-243-539-2
; Sequence 2, Application US/09243539
; Patent No. 6130203
; GENERAL INFORMATION:
; APPLICANT: VOORBERG, Johannes J.
; TITLE OF INVENTION: HYBRID PROTEINS WITH MODIFIED ACTIVITY
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```



SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/243.539  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/558,107  
FILING DATE: 13-NOV-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: ISACSON, John P.  
REGISTRATION NUMBER: 33,715  
REFERENCE/DOCKET NUMBER: 30472/212  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1661 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-243-539-2

Query Match 96.8%; Score 7616; DB 4; Length 1661;  
Best Local Similarity 87.0%; Pred. No. 0;  
Matches 1450; Conservative 5; Mismatches 10; Indels 202; Gaps 4;

QY 1 MEIETSTCFPLCLRFCSATRRYLLGAVELSDMYQSDLGELPDARPPRPVPSFPFN 60  
DB 1 MEIETSTCFPLCLRFCSATRRYLLGAVELSDMYQSDLGELPDARPPRPVPSFPFN 60  
QY 61 TSVYVKKTLFEFFYTHLFIKAPRPMGLGPTQAEVDTVTYTLKNMASHVPSLHAY 120  
DB 61 TSVYVKKTLFEFFYTHLFIKAPRPMGLGPTQAEVDTVTYTLKNMASHVPSLHAY 120  
QY 121 GVSYKASGAEYDQTSQREKEDKVPFGSGSHYVMQVLKENGPMASDPLCTLYSYLSH 180  
DB 121 GVSYKASGAEYDQTSQREKEDKVPFGSGSHYVMQVLKENGPMASDPLCTLYSYLSH 180  
QY 181 VDLVKDLNSGLIGALLVREGSLAKEKTQTLHKFTLLFAVFEDEGKSWHSETKNSLMODRD 240  
DB 181 VDLVKDLNSGLIGALLVREGSLAKEKTQTLHKFTLLFAVFEDEGKSWHSETKNSLMODRD 240  
QY 241 AASAPAMKMTVNGYVRSILPGLIGCHRSYVYHVMGTTPREYHSTFLEGHFPLVANH 300  
DB 241 AASAPAMKMTVNGYVRSILPGLIGCHRSYVYHVMGTTPREYHSTFLEGHFPLVANH 300  
QY 301 ROASLEISPIFLTAQTLMDLGOFLFCHTSSHODGMEAYVYVDSCPREPOLRMKNNE 360  
DB 301 ROASLEISPIFLTAQTLMDLGOFLFCHTSSHODGMEAYVYVDSCPREPOLRMKNNE 360  
QY 361 EAEYDDDLTSEMDVYVREDDNSPSFTQIRSVAKKHFKTVMWYIAAEEEDMDYAPLVLA 420  
DB 361 EAEYDDDLTSEMDVYVREDDNSPSFTQIRSVAKKHFKTVMWYIAAEEEDMDYAPLVLA 420  
QY 421 PDDRSYKSYLNNGPORGKRYKKYVREMAAYDEFEKPEALQOHSSGITLPLLYGEVDTL 480  
DB 421 PDDRSYKSYLNNGPORGKRYKKYVREMAAYDEFEKPEALQOHSSGITLPLLYGEVDTL 480  
QY 481 LIFKNOASRPYNTYPHGTDVPRPLYSRRLPKGVKHLKDFPLFGEIFKYKWTYVEDGP 540  
DB 481 LIFKNOASRPYNTYPHGTDVPRPLYSRRLPKGVKHLKDFPLFGEIFKYKWTYVEDGP 540  
QY 541 TKSPDRCLTRYSSFVNNERDLASGLIGPLLCYKESVDQSGNDIMSDKRVNILLFSVDE 600  
DB 541 TKSPDRCLTRYSSFVNNERDLASGLIGPLLCYKESVDQSGNDIMSDKRVNILLFSVDE 600  
QY 601 NRSWLTENIORFLPNAGVQLEDEFOASIMHSINGYVDSIQLSVCLHEVAYWYILS 660  
DB 601 NRSWLTENIORFLPNAGVQLEDEFOASIMHSINGYVDSIQLSVCLHEVAYWYILS 660

QY 661 IGAQTFDLVSFESGTYFEKHKMYEDTLTLFPFSGETVPMENPGLWILGCHNSDFNRNG 720  
DB 661 IGAQTFDLVSFESGTYFEKHKMYEDTLTLFPFSGETVPMENPGLWILGCHNSDFNRNG 720  
QY 721 MTALIKVSSC---DKNTGDY-----EDSYEDISAYLLSKNNALEPRFSQNSRHPS 769  
DB 721 MTALIKVSSCIPGEDEDDYLDLEKIFSEDDYIDI-----VSLLEPRFSQNSRHPS 774  
QY 770 TRQKQFNT----- 778  
DB 775 TRQKQFNTIIPENDIEKTPWFARHPMPKIQNVSSDMLMLRQSPTPHGLSLDLQE 834  
QY 779 ----- 778  
DB 835 AKYETFSDDPSPGALIDNSNLSMETHFRPOLHSGDMVFTPEBSGLQLRLNEKLCGTTADPL 894  
QY 779 ----- 778  
DB 895 AMDNHGTOIPKEEKKSQEKSPKTAFAKKKDTILSLNACESNHAIAINEGONKPEIEVT 954  
QY 779 ----- 778  
DB 955 WAKQRTERLCSQNPVYLKRQREITRTTLOSDQDEIDYDITISVEMKKEFDIYDEDEN 1014  
QY 825 QSPRSFOKTRHYFLAAVERLMDYGMSSSPHYLNRRAQSGSVPOFKKVPQETDGSFTQ 884  
DB 1015 QSPRSFOKTRHYFLAAVERLMDYGMSSSPHYLNRRAQSGSVPOFKKVPQETDGSFTQ 1074  
QY 885 PLYGELNEHLGLLPYIRAEVEDNIMVTFRNQASRPYSLSLSYEEDROGAEPKRN 944  
DB 1075 PLYGELNEHLGLLPYIRAEVEDNIMVTFRNQASRPYSLSLSYEEDROGAEPKRN 1134  
QY 945 FVKPNEKTFYFMKVOHNAAPKDEFCDAKAAVFSVDVLEKRYHSGLLGPLVCHTNTLNP 1004  
DB 1135 FVKPNEKTFYFMKVOHNAAPKDEFCDAKAAVFSVDVLEKRYHSGLLGPLVCHTNTLNP 1194  
QY 1005 AHGROVTVQEFALFTTFDEKSWYFTEENMRNCRAPNIMEPTPEKENYRFAINGYI 1064  
DB 1195 AHGROVTVQEFALFTTFDEKSWYFTEENMRNCRAPNIMEPTPEKENYRFAINGYI 1254  
QY 1255 MDTLPLGLVMAQDQRIRWYLLSMGSENENIHSIFSGHVFYVKKREYKMAVLYNLPVGFET 1314  
DB 1065 MDTLPLGLVMAQDQRIRWYLLSMGSENENIHSIFSGHVFYVKKREYKMAVLYNLPVGFET 1124  
QY 1125 VEMLPKRAQIRVRCCLIGEHHLAAGMSTFLVYSNKCQTPGLMASGHIRDQITASGOYQ 1184  
DB 1315 VEMLPKRAQIRVRCCLIGEHHLAAGMSTFLVYSNKCQTPGLMASGHIRDQITASGOYQ 1374  
QY 1185 WAPKLARLHSGSINAMSTKEPFMWIKYVDLAPMIHIGITQGAROKFESSLYISQFTIMY 1244  
DB 1375 WAPKLARLHSGSINAMSTKEPFMWIKYVDLAPMIHIGITQGAROKFESSLYISQFTIMY 1434  
QY 1245 SLDGKKMOTYRGNSTGTLMVFFGNVDSGIKHNIFNPPIIARYIRLPHYHSINSTLRME 1304  
DB 1435 SLDGKKMOTYRGNSTGTLMVFFGNVDSGIKHNIFNPPIIARYIRLPHYHSINSTLRME 1494  
QY 1305 LMGCDLNSCNPPLGMEKASIDAOITASSYFTNNFAWWSKAPRLHIOGSSNMAPRVYNN 1364  
DB 1495 LMGCDLNSCNPPLGMEKASIDAOITASSYFTNNFAWWSKAPRLHIOGSSNMAPRVYNN 1554  
QY 1365 PKEWLQVDFOKTMKVTGVTQGVKSLTSMYVKEFLISSQODGHOVTLFPONGKVKVFOG 1424  
DB 1555 PKEWLQVDFOKTMKVTGVTQGVKSLTSMYVKEFLISSQODGHOVTLFPONGKVKVFOG 1614  
QY 1425 NODSFTPVNSLDPILLTRYLRIPHQSVHQAIALRMEVLCCEADLY 1471  
DB 1615 NODSFTPVNSLDPILLTRYLRIPHQSVHQAIALRMEVLCCEADLY 1661

RESULT 5  
US-08-121-202-2  
; Sequence 2, Application US/08121202  
; Patent No. 5563045







```

QY 779 ----- 778
Db 1561 GKVPFLRVATESSAKTPSKLDPFLAMDNHYGTQIPKEEMKSOEKSPKTAFFKKTDTLTL 1620
QY 779 ----- -PVLAKRIQRETTTTLQSDQEE 800
Db 1621 MACESNHAIAINEGONKPEIEVTWAKOGRTERLCSQNPVPLKRHOQRETTTLQSDQEE 1680
QY 801 IDYDDITSEYEMKKEPDIDYDEDEENOSPRSFQKTRHYFIAAVERLMDYMSSPHYLRNR 860
Db 1681 IDYDDITSEYEMKKEPDIDYDEDEENOSPRSFQKTRHYFIAAVERLMDYMSSPHYLRNR 1740
QY 861 AOSGSVPOEKVYVFOETDGSFTQPLYRGELNEHLGLGPLYRAEVEDNIMVTFRNQASR 920
Db 1741 AOSGSVPOEKVYVFOETDGSFTQPLYRGELNEHLGLGPLYRAEVEDNIMVTFRNQASR 1800
QY 921 PVSFYSLSLSTYEDDQOGAPRKNFVKPNETKTYFWKVOHNAAPTKDEDFCKAMAYFSVY 980
Db 1801 PVSFYSLSLSTYEDDQOGAPRKNFVKPNETKTYFWKVOHNAAPTKDEDFCKAMAYFSVY 1860
QY 981 DLEKDVHSLIGPLLYCHNTNINPAHGRQVTVQEFALFTTIDETKSWYFTEEMERNCA 1040
Db 1861 DLEKDVHSLIGPLLYCHNTNINPAHGRQVTVQEFALFTTIDETKSWYFTEEMERNCA 1920
QY 1041 PCNIQMEDPFEKENYRFHAINCYIMDTPLGLVMAODQIRIMYLLSGNSNENHSHIFSGH 1100
Db 1921 PCNIQMEDPFEKENYRFHAINCYIMDTPLGLVMAODQIRIMYLLSGNSNENHSHIFSGH 1980
QY 1101 VETVRRKEEYKALNLYPGVEFETVEMLPKSKAGIMRVECLTGEHLHAGNSTLFLVYSNKC 1160
Db 1981 VETVRRKEEYKALNLYPGVEFETVEMLPKSKAGIMRVECLTGEHLHAGNSTLFLVYSNKC 2040
QY 1161 QPPLGASGHINDFOITTAGOYGNAPKLARLHYSGSINAMSTKEPESIKYDLAPMTI 1220
Db 2041 QPPLGASGHINDFOITTAGOYGNAPKLARLHYSGSINAMSTKEPESIKYDLAPMTI 2100
QY 1221 HGKKTQAGAKOFSSLYISQFIIMYSLDGKKMOTYRGNSTGLTMVFFGNVDSGIGKHNFN 1280
Db 2101 HGKKTQAGAKOFSSLYISQFIIMYSLDGKKMOTYRGNSTGLTMVFFGNVDSGIGKHNFN 2160
QY 1281 PPIIARIYTLFHTHYXSINSTLMELMGCDLNSCSPMLGMSKASIAIDAOITASSYFTNMPA 1340
Db 2161 PPIIARIYTLFHTHYXSINSTLMELMGCDLNSCSPMLGMSKASIAIDAOITASSYFTNMPA 2220
QY 1341 TWSPSKARLHLDGRSNAMPPOVNNPKEMLOVDFOKTMKVTGYTQGVSLTSMYVKEFL 1400
Db 2221 TWSPSKARLHLDGRSNAMPPOVNNPKEMLOVDFOKTMKVTGYTQGVSLTSMYVKEFL 2280
QY 1401 ISSSQDGHQWTLFQNGKVKVYQGNQDSFTPYVNSLDPRLTLRYLRIHQSWVHQIALRM 1460
Db 2281 ISSSQDGHQWTLFQNGKVKVYQGNQDSFTPYVNSLDPRLTLRYLRIHQSWVHQIALRM 2340
QY 1461 EYLGEADOLY 1471
Db 2341 EYLGEADOLY 2351

```

```

; FILLING DATE: 18-NOV-1986
; APPLICATION NUMBER: 868,410
; FILLING DATE: 29-MAY-1986
; SEQ ID NO:1
; LENGTH: 2351
5422260-1

Query Match      94.2% Score 7410; DB 6; Length 2351;
Best Local Similarity 62.5%; Pred No. 0;
Matches 1470; Conservative 0; Mismatches 1; Indels 880; Gaps 1;

QY 1 MOELSTCFELCLLRCSATRRYYLAGVELSMDYMQSDGLGELPYDARPPRPVPSFEPN 60
Db 1 MOELSTCFELCLLRCSATRRYYLAGVELSMDYMQSDGLGELPYDARPPRPVPSFEPN 60
QY 61 TSVVYKKTLEFVEFTVHLFNIAKPRPPWMLGPTIOAEYVDIVVITLTKMAHSPVSLNAV 120
Db 61 TSVVYKKTLEFVEFTVHLFNIAKPRPPWMLGPTIOAEYVDIVVITLTKMAHSPVSLNAV 120
QY 121 GVSYMKASGAEYDDOTSOREKEDDKVPPGSGHTYVMOYLKENGPMASDPLCLTYSLSH 180
Db 121 GVSYMKASGAEYDDOTSOREKEDDKVPPGSGHTYVMOYLKENGPMASDPLCLTYSLSH 180
QY 181 VDLVRDLNSGLIGALLVCREGSLAEKQTLKFTLLFAVDFEGKSWHSETKNSLMODRD 240
Db 181 VDLVRDLNSGLIGALLVCREGSLAEKQTLKFTLLFAVDFEGKSWHSETKNSLMODRD 240
QY 241 AASARAPRMHTVNGVYVNSLPLGLIGCHRKSYVWHVIGMTTPEVHSIFLECHTFLVRNH 300
Db 241 AASARAPRMHTVNGVYVNSLPLGLIGCHRKSYVWHVIGMTTPEVHSIFLECHTFLVRNH 300
QY 301 ROSLEISPIFTPLAQTLLMDLGOFLFCHLSSHQHDGMEAVYKVYDSCPEEBQLMKKNE 360
Db 301 ROSLEISPIFTPLAQTLLMDLGOFLFCHLSSHQHDGMEAVYKVYDSCPEEBQLMKKNE 360
QY 361 EAEDYDDDLDTSEMDVVRFPDDNSPFIQIRSVAKHRTWVHYTAAEEDMDVAPLYLA 420
Db 361 EAEDYDDDLDTSEMDVVRFPDDNSPFIQIRSVAKHRTWVHYTAAEEDMDVAPLYLA 420
QY 421 PDDRSYKSOYLLNNGQORIGRKTKKVRFMAAYTDEFTKTRBAIOHESGILGPLYGEVDTL 480
Db 421 PDDRSYKSOYLLNNGQORIGRKTKKVRFMAAYTDEFTKTRBAIOHESGILGPLYGEVDTL 480
QY 481 LITFEKQASRPVNIYPHGTTDVRPLYSRRLPKGVNHLMDPFLPBEIRKYKTFVVEGDP 540
Db 481 LITFEKQASRPVNIYPHGTTDVRPLYSRRLPKGVNHLMDPFLPBEIRKYKTFVVEGDP 540
QY 541 TKS DPRCLTRYYSFVNNERDLASGLIPLLCYKESYVDORQNOJMSDKRNVILFSVDE 600
Db 541 TKS DPRCLTRYYSFVNNERDLASGLIPLLCYKESYVDORQNOJMSDKRNVILFSVDE 600
QY 601 NRSWLTENIQFLEPNPAGVQLEDEPFOASNTMBSINGVYDSDLSCLLHEVAYWYILS 660
Db 601 NRSWLTENIQFLEPNPAGVQLEDEPFOASNTMBSINGVYDSDLSCLLHEVAYWYILS 660
QY 661 IGAOTDFLSVFPFSGYTFKHKNYEDTLTLPFSGETVPMSPNPGTLWILGCHNSPFRNRG 720
Db 661 IGAOTDFLSVFPFSGYTFKHKNYEDTLTLPFSGETVPMSPNPGTLWILGCHNSPFRNRG 720
QY 721 MTALLKVSCKDNKTDYEDSYEDISAYLLSKNNAIEPRFSQNSRHSSTROKOPNAT-- 778
Db 721 MTALLKVSCKDNKTDYEDSYEDISAYLLSKNNAIEPRFSQNSRHSSTROKOPNAT-- 778
QY 779 ----- 778
Db 781 PENDIEKTDPMFAHRTPMPKIQNVSSDMLMLROSPTPHGLSLSDLOEAKYTFSDPS 840
QY 779 ----- 778
Db 841 PGADISNNSLSMHTFRQLHHSQDMVFTPSGDLRLNEKLGTIAATLELKKLDFKVSST 900
QY 779 ----- 778

```



```

Db 901 SNNLISTIPSDNLAAGTNTSSLGPPSMVHYHDSQDITLLEGGKSSPLSESGPLSLEE 960
Qy 779 ----- 778
Db 961 NNDKSLLESGILNDSQSSGKNVSTESGRLEFKGRAGHGPALLTKDNALFKVISILKTN 1020
Qy 779 ----- 778
Db 1021 KTSNNSATNRKTHIDGPSLLIENSPSWONILESDTEFKKVTPLIHDRMLDKNATALRL 1080
Qy 779 ----- 778
Db 1081 NMSMKTTSKKNMENVQOKKEGPIPDAONPDMSFFKMLFLPESARMIQTHGKNSLNSG 1140
Qy 779 ----- 778
Db 1141 QGSPRKQVLSGLPEKSVESQONLSEKNKYVGGEFTKDVGLKEMVPPSSRNLETLNLDN 1200
Qy 779 ----- 778
Db 1201 LHENNTHNOEKKIOEIEKEKTLIGENVVLPQIHVTGKNEMKMLFLSTRQNVESYE 1260
Qy 779 ----- 778
Db 1261 GAYAPVLODFRSLNDSTNFTKKNHAFHSKKEEENLEGLGNQTKQIVERYACTTRISPMT 1320
Qy 779 ----- 778
Db 1321 SQONVNTQSRKALQOFRLPLETELEKRIYDDTSTQMSKNKHLTPSTLQIDYNEKE 1380
Qy 779 ----- 778
Db 1381 KGAITQPSLDCLTRSHSIPQANRSPPIAKVSSPFSIRPIYLRVLFODNSHLPAASY 1440
Qy 779 ----- 778
Db 1441 RKKDSGVOBSSHFLQAGAKNNLSLALILEMTGDQREVSGLSGTSATNSVYKKVENTVLP 1500
Qy 779 ----- 778
Db 1501 KPDLPKTSGKVELLRKVHLYQKDLPTETNSGPHGLDVEGSLLOCTGCAIKWNEANRP 1560
Qy 779 ----- 778
Db 1561 GKVELRVATESSAKTPSKLLDPLAMDNHYGTQIPKEEMKSOEKSEPTAFKKOTLILSL 1620
Qy 779 ----- -PVLKRHQREITRTTLOSDOE 800
Db 1621 NACESNHAIAINEGONKPEIEVYAKQGRTERLCSQNPVPLKRHQREITRTTLOSDOE 1680
Qy 801 IDYDDTISVEMKEDFDIYDEDENSPSPFOKTRHYFAAVERLMDYGMSSSPHYLRNR 860
Db 1681 IDYDDTISVEMKEDFDIYDEDENSPSPFOKTRHYFAAVERLMDYGMSSSPHYLRNR 1740
Qy 861 AOSGVPQPKVYFOEFDGFTQPLYRGELNENHGLGCPYRAEVEDNIMVTFRNQASR 920
Db 1741 AOSGVPQPKVYFOEFDGFTQPLYRGELNENHGLGCPYRAEVEDNIMVTFRNQASR 1800
Qy 921 PYSFSSSLISYEEDROGAEPRKNFVKPNETKTYEMKVYOHNAHAPRTDEEDCKAMAEFSV 980
Db 1801 PYSFSSSLISYEEDROGAEPRKNFVKPNETKTYEMKVYOHNAHAPRTDEEDCKAMAEFSV 1860
Qy 981 DLEKDVHSGILGPLLVCHTNTLNPAGROVYQEFALFEITDEFKSWYFTENMEHNCRA 1040
Db 1861 DLEKDVHSGILGPLLVCHTNTLNPAGROVYQEFALFEITDEFKSWYFTENMEHNCRA 1920
Qy 1041 PCNTOMEDPTEKENTRFAHNGYIMDTPLGLVYADQRTIRWYLLSGNSNENHSHISFSGH 1100
Db 1921 PCNTOMEDPTEKENTRFAHNGYIMDTPLGLVYADQRTIRWYLLSGNSNENHSHISFSGH 1980
Qy 1101 VFTYRKKKEEYKALYNLYPGVETVEMLPKSGKIMRVELLIGEHLAGNSTLFLVYSNKC 1160

```

```

Db 1981 VFTYRKKKEEYKALYNLYPGVETVEMLPKSGKIMRVELLIGEHLAGNSTLFLVYSNKC 2040
Qy 1161 QPPLGASGHIRDFQITASGOYGOMAPKARLHYSGSINAWSTKEPFEMIKVDLAPMI 1220
Db 2041 QPPLGASGHIRDFQITASGOYGOMAPKARLHYSGSINAWSTKEPFEMIKVDLAPMI 2100
Qy 1221 HGKTOGAKOKSSSLYSOPIIMYSLDGKKWOTYRGNSTGLMVFEGVNDSSGKIHNIEN 1280
Db 2101 HGKTOGAKOKSSSLYSOPIIMYSLDGKKWOTYRGNSTGLMVFEGVNDSSGKIHNIEN 2160
Qy 1281 PPIIARYIRLPHYHSIRSTLAMELMGCDLNSCSPMLGMSKSAISDAOITASSYFTNMEA 1340
Db 2161 PPIIARYIRLPHYHSIRSTLAMELMGCDLNSCSPMLGMSKSAISDAOITASSYFTNMEA 2220
Qy 1341 TWSPSKARLHLOGRSNAMPPOVNNPEKEMLOVDFOKTMKYTGTTQGVKSLTSMYKEFL 1400
Db 2221 TWSPSKARLHLOGRSNAMPPOVNNPEKEMLOVDFOKTMKYTGTTQGVKSLTSMYKEFL 2280
Qy 1401 ISSODGHQWTLFPOGKVKVFOGNDSFTPVVNSLDPLRLRYLRHPOKWHQIALRM 1460
Db 2281 ISSODGHQWTLFPOGKVKVFOGNDSFTPVVNSLDPLRLRYLRHPOKWHQIALRM 2340
Qy 1461 EYLGCCEADPLY 1471
Db 2341 EYLGCCEADPLY 2351

```

```

RESULT 7
5171844-2
; Patent No. 5171844
; APPLICANT: VAN OOEYEN, ALBERT J.J.; PANNEKOEK, HANS; VERBEET,
; MARTINUS P.; VAN LEEUW, ROBERT W.
; TITLE OF INVENTION: PROTEINS WITH FACTOR VIII ACTIVITY
; PROCESS FOR THEIR PREPARATION USING GENETICALLY-ENGINEERED CELLS
; AND PHARMACEUTICAL COMPOSITIONS CONTAINING THEM
; NUMBER OF SEQUENCES: 12
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/205,226
; FILING DATE: 10-JUN-1988
; SEQ ID NO: 2:
; LENGTH: 2351
5171844-2

```

```

Query Match 94.1%; Score 7406; DB 6; Length 2351;
Best local similarity 62.5%; Pred. No. 0;
Matches 1469; Conservative 0; Mismatches 2; Indels 880; Gaps 1;

Qy 1 MOTELSTCFELCLLRFCFSATRRYLLGAVELSDMYQSDGLBEPVDARPPRRVRSPPFN 60
Db 1 MOTELSTCFELCLLRFCFSATRRYLLGAVELSDMYQSDGLBEPVDARPPRRVRSPPFN 60
Qy 61 TSVYVKKTLFEFTYHLFPIAKPRPMGLGPTQAEVYDTVYTLKNMASHVSLHAV 120
Db 61 TSVYVKKTLFEFTYHLFPIAKPRPMGLGPTQAEVYDTVYTLKNMASHVSLHAV 120
Qy 121 GVSYKASGAEYDDQTSQREKEDKVPFGSGHTYVMQVLKENGEMASDPLCLTYSLSH 180
Db 121 GVSYKASGAEYDDQTSQREKEDKVPFGSGHTYVMQVLKENGEMASDPLCLTYSLSH 180
Qy 181 VDLVYDNLNGLIGALLVCEGSLAKEKOTLHKLFLFAVDEGKSWSEFRKNSLMODRD 240
Db 181 VDLVYDNLNGLIGALLVCEGSLAKEKOTLHKLFLFAVDEGKSWSEFRKNSLMODRD 240
Qy 241 AASARAMPKMHVNVNYSGLGICGRKSYVMHVTGMGTPEVHSFELGHTFIVRNH 300
Db 241 AASARAMPKMHVNVNYSGLGICGRKSYVMHVTGMGTPEVHSFELGHTFIVRNH 300
Qy 301 ROASLEISPTFLTAQTLMDGQFLFCCHISSHQDMEAYVYKDSCEPPEQRLRMKNE 360
Db 301 ROASLEISPTFLTAQTLMDGQFLFCCHISSHQDMEAYVYKDSCEPPEQRLRMKNE 360
Qy 361 EAEDYDDDLTDEMDVVRDNDNSPFTQIRSVAKKRPKRWVHYIAEEDMDYAPLVLA 420

```



```

|||||
Db 361 EAEEDDDDLTJSEMDYVRFDDDNSEFIOIRSVAKKHPKTHVHYIAAEEDMDYAPLVLA 420
Oy 421 PDRSRKSSOYLNGPORGKRYKKVRFMAVYDETEKTRREALIOHESGIIIGPLLYGVGDTL 480
Db 421 PDRSRKSSOYLNGPORGKRYKKVRFMAVYDETEKTRREALIOHESGIIIGPLLYGVGDTL 480
Oy 481 LIIFFNOASRPNIYPHGITDVRPLYSRRLPKGVKHLKDFILPEIEIKKYKWTVEDEGP 540
Db 481 LIIFFNOASRPNIYPHGITDVRPLYSRRLPKGVKHLKDFILPEIEIKKYKWTVEDEGP 540
Oy 541 TKSDEPRCLTRYSSFFVNNEBDLASGLIGPLLYCYESVDQRCNOJMSDKRNVILFESVEDE 600
Db 541 TKSDEPRCLTRYSSFFVNNEBDLASGLIGPLLYCYESVDQRCNOJMSDKRNVILFESVEDE 600
Oy 601 NNSWLTENIQOFLPNPAGVOLEDEPEFQASNTMHSINGVDPDSLOLSCVHEVAWYILS 660
Db 601 NNSWLTENIQOFLPNPAGVOLEDEPEFQASNTMHSINGVDPDSLOLSCVHEVAWYILS 660
Oy 661 IGAOTDFLSVFESGYTFKHKMYEDTLFLFPESGETVFMSEMNPGLMILGCHNSDFRNNG 720
Db 661 IGAOTDFLSVFESGYTFKHKMYEDTLFLFPESGETVFMSEMNPGLMILGCHNSDFRNNG 720
Oy 721 MTALLKVVSCOKNTGDYEDSYEDISAYLLSKNNAIEPRSFQONSRRHSTROKOFNAT - 778
Db 721 MTALLKVVSCOKNTGDYEDSYEDISAYLLSKNNAIEPRSFQONSRRHSTROKOFNAT - 780
Oy 779 ----- 778
Db 781 PENDIEKTDPMFAHRTPMFKIONVSSSDDLMLMLROSPTPHGISLSLDQBAKYETFSDDPS 840
Oy 779 ----- 778
Db 841 PGALDSNNSLSBMTFRPOLHHSQDMVFTPESGLOLRINEKLGTTAATELKKLDPKVSST 900
Oy 779 ----- 778
Db 901 SNMLISTIPDNLAGTNDTSSISGPPSMVHYDQLODTFLFGKSSPLTESGGLSLEE 960
Oy 779 ----- 778
Db 961 NNDKSLLESGLMNSQESSMGKNVSTESGRLFKGRRAHPALTTKDNALFKVYSISLKTN 1020
Oy 779 ----- 778
Db 1021 KTSNNSATNRKTHIDGPSLLIENSPVONILIESDTEFKKYTPPLIHDMLMDKATALRL 1080
Oy 779 ----- 778
Db 1081 NHMSNKTSSKMEMVQOKKEGPIPPDAQNPDMSPFKMLFLPESARWIOPTHGKNSLNSG 1140
Oy 779 ----- 778
Db 1141 QGSPKQVLVSLGPEKSVEGONFLSEKNKVYVVGKEFTKDVGLKEMVFPSSRNLFLTNLDN 1200
Oy 779 ----- 778
Db 1201 LHENNTHOEKKIOEIEIKKETLIOENVVLOIHTVTGKKNFMKNLFLSTRONVGSYD 1260
Oy 779 ----- 778
Db 1261 GAYAPVLQDFRSLNDSTNRKTHKHTAHFSKGBEENLEGLIGNOTKOIVEKACTTRISPNT 1320
Oy 779 ----- 778
Db 1321 SOONFVQSRKRAIKOFLPLEETELEKRIIVDSTQMSKNMKHLNPSLTLOIDYNEKE 1380
Oy 779 ----- 778
Db 1381 KGATQSPSLSDCLTRSHSIPQANKSPPLIAKVSSPSTIRPIYLTIVLFDQDNSSHLPAASY 1440
Oy 779 ----- 778

```

```

Db 1441 RKKDSGVOESSHFILOGAKKNNLSLAILTLEMTGQREVSGTGSATNSVTKKVENTVLP 1500
Oy 779 ----- 778
Db 1501 KPDLFKTSGKVELLPKHVHIYOKDLFPTETSNGPSGLDLVBSGLQTEGAIKMNEANRP 1560
Oy 779 ----- 778
Db 1561 GKVPFLRATESSAKTPSKLDDPLAMDHNHYGTQIPKEEMKSOEKSPEKTAFKKKDTIISL 1620
Oy 779 ----- 1620
Db 1621 NACESNHAIAAINESQNKPEIEVTWAKOGRTERLCQNPVYLKRHOREITRTTLOSDOE 1680
Oy 801 IDYDITISVEKKKEDEFDIYDEDENSPSPQKTPHYTIAVERIMDYGSSPVLNRR 860
Db 1661 IDYDITISVEKKKEDEFDIYDEDENSPSPQKTPHYTIAVERIMDYGSSPVLNRR 1740
Oy 861 AOSGSVPQFKKVPFOEFTDGSFTQPLRYGELNEHLGLGPYIRAEVENINIMVTFRNQASR 920
Db 1741 AOSGSVPQFKKVPFOEFTDGSFTQPLRYGELNEHLGLGPYIRAEVENINIMVTFRNQASR 1800
Oy 921 PYSFYSSLISYEEDQROGAEPKKNFVKDNEKTYFKVQVHMAPTKDEPDKAMAYFSDV 980
Db 1801 PYSFYSSLISYEEDQROGAEPKKNFVKDNEKTYFKVQVHMAPTKDEPDKAMAYFSDV 1860
Oy 981 DLEKDVHSLIGPLLYVCHTNLTPAHGROYVVOEALFETTFDEKSKYFTENMRNCRA 1040
Db 1861 DLEKDVHSLIGPLLYVCHTNLTPAHGROYVVOEALFETTFDEKSKYFTENMRNCRA 1920
Oy 1041 PCNIQMEDPTEKENYRFAINNGYIMDTLPGVMAQDORIRWYLSMGSNENIHSHFSGH 1100
Db 1921 PCNIQMEDPTEKENYRFAINNGYIMDTLPGVMAQDORIRWYLSMGSNENIHSHFSGH 1980
Oy 1101 VFTVARKKEEYKALYNLYPVFEVEMLPKSAQIWRVECLIGEHLHAQMSTLFLVYSNKC 1160
Db 1981 VFTVARKKEEYKALYNLYPVFEVEMLPKSAQIWRVECLIGEHLHAQMSTLFLVYSNKC 2040
Oy 1161 QTPGLMASGHTRDQITASGOYGQWAPKLARLHYSGSINASTKPEFMIYVDLAPMI 1220
Db 2041 QTPGLMASGHTRDQITASGOYGQWAPKLARLHYSGSINASTKPEFMIYVDLAPMI 2100
Oy 1221 HGIKTOGAROKFESSLYIQFIIMYSLDGKKMOTYRGNSGTGLMVFNGVDSGKIHNFEN 1280
Db 2101 HGIKTOGAROKFESSLYIQFIIMYSLDGKKMOTYRGNSGTGLMVFNGVDSGKIHNFEN 2160
Oy 1281 PPIIARYTRLPHTHYSIRSTLRMELMGCDLNSCSMPLGMESKASISDAQITASSYFTNMFA 1340
Db 2161 PPIIARYTRLPHTHYSIRSTLRMELMGCDLNSCSMPLGMESKASISDAQITASSYFTNMFA 2220
Oy 1341 TWSPEKARLHLQGRSNARPOVNNKKEWLOYDFOKTMTKVTGVTTOGVVSLTSMYVKEFL 1400
Db 2221 TWSPEKARLHLQGRSNARPOVNNKKEWLOYDFOKTMTKVTGVTTOGVVSLTSMYVKEFL 2280
Oy 1401 ISSSODGHQWTLFQNGVYKFKFGQMODSFTPVVNSLDPLLTRYLRIRIPQSVHIOIALRM 1460
Db 2281 ISSSODGHQWTLFQNGVYKFKFGQMODSFTPVVNSLDPLLTRYLRIRIPQSVHIOIALRM 2340
Oy 1461 EYLGECAODLY 1471
Db 2341 EYLGECAODLY 2351

```

```

RESULT 8
US-08-366-851A-2
: Sequence 2, Application US/08366851A
: Patent No. 5681746
: GENERAL INFORMATION:
: APPLICANT: Bodner, Mordechai
: APPLICANT: De Polo, Nicolas J.
: APPLICANT: Hsu, David Chi-Tang
: APPLICANT: Chang, Steven
: TITLE OF INVENTION: Retroviral Delivery of Full Length Factor VIII

```



NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Viagen, Inc.  
STREET: 11055 Roselle Street  
CITY: San Diego  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 92121  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/366,851A  
FILING DATE:  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Chambers, Daniel M.  
REGISTRATION NUMBER: 34,561  
REFERENCE/DOCKET NUMBER: 930049.438  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 452-1288  
TELEFAX: (619) 452-2616  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2351 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-366-851A-2

Query Match 94.1%; Score 7403; DB 1; Length 2351;  
Best Local Similarity 62.5%; Pred. No. 0;  
Matches 1469; Conservative 0; Mismatches 2; Indels 880; Gaps 1;  
QY 1 MOELSTFCLCLRPSCATRRYYIGAVELSMYQSDGLPVDARPPRYPKSFPFN 60  
DB 1 MOELSTCFCLCLRFCSATRRYYIGAVELSMYQSDGLPVDARPPRYPKSFPFN 60  
QY 61 TSVYVKTLFVEFTVHLFIAPRPPWMLGPTQAEVYDVTVTLKMMASHPVSLHAV 120  
DB 61 TSVYVKTLFVEFTVHLFIAPRPPWMLGPTQAEVYDVTVTLKMMASHPVSLHAV 120  
QY 121 GVSYKASEGAEYDQTSQREKEDDKVFPGSHTYWQVLYKENGPMASDPLCLTYSLSH 180  
DB 121 GVSYKASEGAEYDQTSQREKEDDKVFPGSHTYWQVLYKENGPMASDPLCLTYSLSH 180  
QY 121 GVSYKASEGAEYDQTSQREKEDDKVFPGSHTYWQVLYKENGPMASDPLCLTYSLSH 180  
DB 121 GVSYKASEGAEYDQTSQREKEDDKVFPGSHTYWQVLYKENGPMASDPLCLTYSLSH 180  
QY 181 VDLVYKNSGLIGALLVCREGSLAKETQTLHKFILLFAVDEGKSMHSETKNSLMQDRD 240  
DB 181 VDLVYKNSGLIGALLVCREGSLAKETQTLHKFILLFAVDEGKSMHSETKNSLMQDRD 240  
QY 241 AASARAPKMTNVNGVYVRSRLPGILGCHRSYVYHIGGTPPEVHISILEGHTFLVRNH 300  
DB 241 AASARAPKMTNVNGVYVRSRLPGILGCHRSYVYHIGGTPPEVHISILEGHTFLVRNH 300  
QY 301 ROASLEISPIITFLAQOTLLMDGQFLFCHISSHQHDMGEAVYKVDSCPEEPOLRMKNNE 360  
DB 301 ROASLEISPIITFLAQOTLLMDGQFLFCHISSHQHDMGEAVYKVDSCPEEPOLRMKNNE 360  
QY 361 EAEYDDELTDSEMDVYVREDNNSPSFIQIRSVAKKHPTWYHYIAEEDMDYADLVLA 420  
DB 361 EAEYDDELTDSEMDVYVREDNNSPSFIQIRSVAKKHPTWYHYIAEEDMDYADLVLA 420  
QY 421 PDDSYKSOYLNNGPQRIGRKKYKRYREMAVYDEFTREAIQHESGILLPLLYGEGDPL 480  
DB 421 PDDSYKSOYLNNGPQRIGRKKYKRYREMAVYDEFTREAIQHESGILLPLLYGEGDPL 480  
QY 481 LIIFKQASRPYNYLPHGITDVRPLYSRLLPKGVKHLKDFPLPGEIFKYKMTVYEDGP 540  
DB 481 LIIFKQASRPYNYLPHGITDVRPLYSRLLPKGVKHLKDFPLPGEIFKYKMTVYEDGP 540

QY 541 TKSDFRCLTRYISSFVNMERDLASGLIGPLLICYESVDQRGNOIMSKRNVIYFSFDE 600  
DB 541 TKSDFRCLTRYISSFVNMERDLASGLIGPLLICYESVDQRGNOIMSKRNVIYFSFDE 600  
QY 601 NRSWYLTENIORFLPAPAGVQLEDPEFOASNIMHSNGVYPSIOLSVCLHEVAWYLLS 660  
DB 601 NRSWYLTENIORFLPAPAGVQLEDPEFOASNIMHSNGVYPSIOLSVCLHEVAWYLLS 660  
QY 661 IGAOTDFLSVFSGYTFKKRMVYEDTLTFPESGETVPMSENPGLMTLIGCHNSDFNRNG 720  
DB 661 IGAOTDFLSVFSGYTFKKRMVYEDTLTFPESGETVPMSENPGLMTLIGCHNSDFNRNG 720  
QY 721 MTALLKVVSCDKNTGDIYDSTEDISAVYLLSKNNAIEPPRSQNSRHPSTROKOPNAT 778  
DB 721 MTALLKVVSCDKNTGDIYDSTEDISAVYLLSKNNAIEPPRSQNSRHPSTROKOPNAT 778  
QY 779 ----- 778  
DB 781 PENDLEKTPWFARHTPMKIQNVSSDLMILLRQSPTRPHGLSLDLOAKYETPSDDPS 840  
QY 779 ----- 778  
DB 841 PGALDSNNSLSEMTHRPQLHSHGDMVFTPESGLQRLNEKLGTTAATELKKLDFKVSST 900  
QY 779 ----- 778  
DB 901 SNMLSTIPSDNLAGTDNTSSLGPPMPVHYDQLDLTFEGKSSPLTESGGLSLEE 960  
QY 779 ----- 778  
DB 961 NNDKSLLESGLMNSQESSGKKNVSTESGRLEFKRAHGPALLIKDNLKRVYSISLKTN 1020  
QY 779 ----- 778  
DB 1021 KTSNNSATNRKTHIDGPSLLIENSPVMONIESDTERKKVYPLIHDMLMDKATLRL 1080  
QY 779 ----- 778  
DB 1081 NMSNKTTSKMNEMVQAKKEGRIPPDAQNPDMSEFKMLFLESARWIORTHGKNSLSNG 1140  
QY 779 ----- 778  
DB 1141 QGSPKQVLSLGEKSVESQGNFLSEKNKYVGKGETYDVLKENVFPFSRNLFLTNLDN 1200  
QY 779 ----- 778  
DB 1201 LHENNTNHEKKIQEIEKKEKTLIDENVYLPQIHVTGKNFMKMLFLLSTRQWEGSYD 1260  
QY 779 ----- 778  
DB 1261 GAVAPYLODFRSLNDSTNRTKKTHTAFSKGKEENLEGLGQTKQIVERKACTTRISPNT 1320  
QY 779 ----- 778  
DB 1321 SQONFVQSRKALQOFLRPLETELEKRIIVDDTSTQMSKMKHLPTSTLTQIDYNEKE 1380  
QY 779 ----- 778  
DB 1381 KGALQPSLSDCLTRSHSIPQANRSPPLIAKVSFSPRIYLRVLEQDNNSHLPASNY 1440  
QY 779 ----- 778  
DB 1441 RKKDSGVQESSHFLQAKKNLSLAILLEMTGDQREVSLGTSATNSVYKKVENTVLP 1500  
QY 779 ----- 778  
DB 1501 KPDLPKTSKVELLPKVHITYQKDLFPTETSNGPSGLDLVBGSLLOGEGALKKNEANRP 1560  
QY 779 ----- 778  
DB 1561 GKVPFLRVATESSAKTPSKLLDPLAMDNIHYGQIQKEEMKSGEKSPEKTAFFKKDTILSL 1620  
QY 779 ----- 800  
-----PVLAKRHQREITRTTIQSDQE



```

Db 1621 NACESNHAIAINEGONKPELEVWAKOGRTERLCSQNPVYLKRHOREITRTTQSDDE 1680
Qy 801 IDYDPTISEVEMKKEPDYIDEENOSPRSFOKKTHTYFIAAVERLMDYGMSSPVVLNR 860
Db 1661 IDYDPTISEVEMKKEPDYIDEENOSPRSFOKKTHTYFIAAVERLMDYGMSSPVVLNR 1740
Qy 861 AOSGSVPQFKVVOEFDFDSFTOPLYRGELNEHLGGLPYIRAEVENINWTFPNOASR 920
Db 1741 AOSGSVPQFKVVOEFDFDSFTOPLYRGELNEHLGGLPYIRAEVENINWTFPNOASR 1800
Qy 921 PYSFYSLSIYSEEDOROGAEBRKNFVKPNETKTYFWKVQHMAAPTDEFDCKAMAYFSDV 980
Db 1801 PYSFYSLSIYSEEDOROGAEBRKNFVKPNETKTYFWKVQHMAAPTDEFDCKAMAYFSDV 1860
Qy 981 DLEKDVHSGLLGPLLVCTNTNLNPAHQVVOEFALEFTTJFDETKSWYFTENNERCRA 1040
Db 1861 DLEKDVHSGLLGPLLVCTNTNLNPAHQVVOEFALEFTTJFDETKSWYFTENNERCRA 1920
Qy 1041 PCNTOMEDPTREKYNRFAINGYIMDTLPGLVMAODRIRNYLLSMGSNENIHSHESGH 1100
Db 1921 PCNTOMEDPTREKYNRFAINGYIMDTLPGLVMAODRIRNYLLSMGSNENIHSHESGH 1980
Qy 1101 VETVARKKEEYKMAALYNLYPGVEFTEVEMLPKAGIMRVECLIGEHLHAGMSTLFLVYSNKC 1160
Db 1981 VETVARKKEEYKMAALYNLYPGVEFTEVEMLPKAGIMRVECLIGEHLHAGMSTLFLVYSNKC 2040
Qy 1161 QTPILGMAHGIRDPQITASGOYGOMAPKLARLHSGSINAMSTEPEFSWIKVDLAPMI 1220
Db 2041 QTPILGMAHGIRDPQITASGOYGOMAPKLARLHSGSINAMSTEPEFSWIKVDLAPMI 2100
Qy 1221 HGITOGAROKFESSLYISQFIMYSLDCKKMOYRGNSTGLMVPFGVNDSSGIKHNEN 1280
Db 2101 HGITOGAROKFESSLYISQFIMYSLDCKKMOYRGNSTGLMVPFGVNDSSGIKHNEN 2160
Qy 1281 PPIIARYIRLHPHTYISINSTLRLMELMGCDLNSCNPILGMEKASISDAQITASSYFTNMA 1340
Db 2161 PPIIARYIRLHPHTYISINSTLRLMELMGCDLNSCNPILGMEKASISDAQITASSYFTNMA 2220
Qy 1341 TWSPSKARLHLGGRSNMARPVNNPKEMLOVDFOKTKMKTGVTTOGVKSLTSMYKEFL 1400
Db 2221 TWSPSKARLHLGGRSNMARPVNNPKEMLOVDFOKTKMKTGVTTOGVKSLTSMYKEFL 2280
Qy 1401 ISSSODGHOFTLFPONGVVKYFQGNOSFTPVNSLOPDLRLRLRHPOSWHOIALRM 1460
Db 2281 ISSSODGHOFTLFPONGVVKYFQGNOSFTPVNSLOPDLRLRLRHPOSWHOIALRM 2340
Qy 1461 EVLGCEAODLY 1471
Db 2341 EVLGCEAODLY 2351

```

```

RESULT 9
US-07-864-004B-4
; Sequence 4, Application US/07864004B
; Patent No. 5364771
; GENERAL INFORMATION:
; APPLICANT: Lollari, John S.
; APPLICANT: Runge, Marschall S.
; TITLE OF INVENTION: Hybrid Human/Porcine Factor VIII
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick & Cody
; STREET: 1100 Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: US
; ZIP: 30309
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25

```

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/864,004B
; FILING DATE: 07 APRIL 1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: EMU106
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-815-6508
; TELEFAX: 404-815-6555
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2332 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapien
; TISSUE TYPE: Liver cDNA sequence
; US-07-864-004B-4

Query Match 93.0%; Score 7315; DB 1; Length 2332;
Best Local Similarity 62.3%; Pred. No. 0;
Matches 1452; Conservative 0; Mismatches 0; Indels 880; Gaps 1;

Qy 20 ATRRYVIGAVELSDMYDSIDGELPVDAKFRPPVKSFPFNTSVYKKTLEVEFVHLFN 79
Db 1 ATRRYVIGAVELSDMYDSIDGELPVDAKFRPPVKSFPFNTSVYKKTLEVEFVHLFN 60
Qy 80 IAKRPPMAGLIGPTIOAEVYDVTYITLKNNASHPVSLHAGVSYWKASEGAEDDQTSQ 139
Db 61 IAKRPPMAGLIGPTIOAEVYDVTYITLKNNASHPVSLHAGVSYWKASEGAEDDQTSQ 120
Qy 140 REKEDDKVFPGSGSHYVQVLKENGPMASDPLCTYSYLSHVLDLVKDNSLIGALLVCR 199
Db 121 REKEDDKVFPGSGSHYVQVLKENGPMASDPLCTYSYLSHVLDLVKDNSLIGALLVCR 180
Qy 200 EGSIAKENTQILHFFILFAVFDGKSMHSETKNSLMQDRDAASARAPKMHNTVNGYNR 259
Db 181 EGSIAKENTQILHFFILFAVFDGKSMHSETKNSLMQDRDAASARAPKMHNTVNGYNR 240
Qy 260 SLPGILGCHRKSYWVHYGMCTPEVHSIFLEGHTFLVRNHRQASLEISPTTFLTAOTLL 319
Db 241 SLPGILGCHRKSYWVHYGMCTPEVHSIFLEGHTFLVRNHRQASLEISPTTFLTAOTLL 300
Qy 320 MDLGOFLFCHISSHQHDMGMAVYKVDSCPEPOLRMKNNEADYDDDLTDEMDVYRF 379
Db 301 MDLGOFLFCHISSHQHDMGMAVYKVDSCPEPOLRMKNNEADYDDDLTDEMDVYRF 360
Qy 380 DDDNSPFIQIRSAKHKPKWVHYIAAEEEDMWYAPLYLAPDRSKYSQVLNNGPQIG 439
Db 361 DDDNSPFIQIRSAKHKPKWVHYIAAEEEDMWYAPLYLAPDRSKYSQVLNNGPQIG 420
Qy 440 RYKRVFMATYDTEFTKREAIQIHESGLIGPLLGEVGDTLILFFKNQASPPNYNYPGI 499
Db 421 RYKRVFMATYDTEFTKREAIQIHESGLIGPLLGEVGDTLILFFKNQASPPNYNYPGI 480
Qy 500 TDVRPLYSRLLPKGVKHLKDEPILPGETFKYKWTVVEDGPTKSDPCLTRYSSFYVME 559
Db 481 TDVRPLYSRLLPKGVKHLKDEPILPGETFKYKWTVVEDGPTKSDPCLTRYSSFYVME 540
Qy 560 RDLASGLIGPLLYCKESVQDQGNQIMSDKRNVLFSVPDNRSWYLTENTQRLPMPAG 619
Db 541 RDLASGLIGPLLYCKESVQDQGNQIMSDKRNVLFSVPDNRSWYLTENTQRLPMPAG 600
Qy 620 VQLEDPFOASNIMHSGVYFDSIQSLVCLHEVAAYWYISIGAQDTFLSVFFSGYTFKH 679
Db 600 VQLEDPFOASNIMHSGVYFDSIQSLVCLHEVAAYWYISIGAQDTFLSVFFSGYTFKH 580

```



Db	601	VQLEPPEQASNIHMSINGVYFDSLQSLVCLHEVAWYILSICAGQDPDLSVEFSGYTFKH	660
Qy	660	KMYVEDTLTLPFGGEYFVMSMEMPGWILIGCHNSDFRNRGMTALLKVSCCKNGDYIE	739
Db	661	KMYEDTLTLPFGGEYFVMSMEMPGWILIGCHNSDFRNRGMTALLKVSCCKNGDYIE	720
Qy	740	DSYEDISAYLLSKNNATEPRFSFONSRRPSPROKOFNAT-----	778
Db	721	DSYEDISAYLLSKNNATEPRFSFONSRRPSPROKOFNATIPENDIEKTDPMFAHRTMP	780
Qy	779	-----	778
Db	781	KIONVSSDLMMLRQSPTRPHGLSLSDLOEAKYETSPDPSGAIDSNNSLSBMTHFQP	840
Qy	779	-----	778
Db	841	LHSGDMVFPPESGLIOLRNEKLGTTAATELKLDPKVSSTSNLSTIPSDNLAAGTDN	900
Qy	779	-----	778
Db	901	TSSLGPPSMVHYDSQDRTLFGKXSSPLTESGGPLSEENDSKLBSGLMNSOESSW	960
Qy	779	-----	778
Db	961	GKNVSTESGRLFKGRRAHGPALLTKNALFKVYSILKTKNTKSNNASATNRKTHIDPSL	1024
Qy	779	-----	778
Db	1021	LIENSPSYMQLIESDTEFKKVTPLIHDRLMDKNATALLRLNMSNKTTSKRMEMVOOK	1080
Qy	779	-----	778
Db	1081	KEGTPPDANQPMDSFFKMLFLPESARMIQRTHGKNSLNSGQSPSKQLVSLGPEKSEVG	1144
Qy	779	-----	778
Db	1141	QNFLEKXKVVVGGEFTKVDGLKEMFPSSRNLFNLNDNLHENTHNOEKKIQEIEIK	1200
Qy	779	-----	778
Db	1201	KETLIOENVLPQIHTVTGFKNEFKNLFLLSTRONEGSEYGAAYALODFRSLNDSTNR	1260
Qy	779	-----	778
Db	1261	TKKHTAHFSKKGEENLEGLGNQTKQIVERYACTTRISPNTSOQNFVQRRSKRALKQERL	1320
Qy	779	-----	778
Db	1321	PLEETELEKRIIVDDTSTQMSKNNKHLTPSTLVOIDYNEKCAITQSPSLDCLTRSHSI	1380
Qy	779	-----	778
Db	1381	QOANRSPPLIAKVSPSPSIRPIYLTVRLFQDNSSHLPASRYRKKDGVQESSHFLQAKK	1440
Qy	779	-----	778
Db	1441	NNLSAILITTEMTGDQREVSLGTSATNSVYRKVENTYLPKPDLPKTSKVYELLPKVHI	1500
Qy	779	-----	778
Db	1501	YOKDLFPETISNGSGHLDVEGSLLOGTEGAIKKNANRPGKVPPLRVATESSAKTFSK	1560
Qy	779	-----	778
Db	1561	LDDPLAMDNHYGTQIPKREMKSQEKSPKTAFFKKDITLSLNCESNHAIAINQGNKP	1624
Qy	779	-----PPVLRKHOREITRTTLOSQDEIDYDDTISEMKKEDFDY	819
Db	1621	EIEVTMAKQGTERLCSQNPVYLKRHRQREITRTTLOSQDEIDYDDTISEMKKEDFDY	1680
Db	1681	DEDENOSPSPOKTTRHYETIAAVERLMDYGSSSPHYLRNRQSGVSPQFKVVFQEFMD	879
Db	1681	DEDENOSPSPOKTTRHYETIAAVERLMDYGSSSPHYLRNRQSGVSPQFKVVFQEFMD	1740

Qy	880	GSEFOPLYRGCLNHNHGLGLCPYIRAEVEDNIMVTFERNKASRPYSYSSLSISEEDORGA	939
Db	1741	GSEFOPLYRGCLNHNHGLGLCPYIRAEVEDNIMVTFERNKASRPYSYSSLSISEEDORGA	1800
Qy	940	EPRRNFPKPNMTEKTYEKKVVOHNAAPTKEFDCKKAMAYFSDVDLEKDVHSGLIGPLVLCNT	999
Db	1801	EPRRNFPKPNMTEKTYEKKVVOHNAAPTKEFDCKKAMAYFSDVDLEKDVHSGLIGPLVLCNT	1860
Qy	1000	NTLNPAHGRVOYVOERPLFTTIDEFTKSWYFTFENMERNCRAPCNIOMEDPTFKENYRPHA	10599
Db	1861	NTLNPAHGRVOYVOERPLFTTIDEFTKSWYFTFENMERNCRAPCNIOMEDPTFKENYRPHA	1920
Qy	1060	INGYINMOTLEGLVWAQOORIRWYLLSMGSENENHSHFEGHVFYVRKKDEEYKMALYNLYP	1119
Db	1921	INGYINMOTLEGLVWAQOORIRWYLLSMGSENENHSHFEGHVFYVRKKDEEYKMALYNLYP	1980
Qy	1120	GVFEFVEMLEPSKAGIMWEVECLLIGEHLHAGNSTLELVYNSKNCOTPLGMASGHIRDFQTAS	11799
Db	1981	GVFEFVEMLEPSKAGIMWEVECLLIGEHLHAGNSTLELVYNSKNCOTPLGMASGHIRDFQTAS	2040
Qy	1180	GQYQOMAPKRLARLHYSSSINAMSTKEPFSWIKYDLAPMITHGITQCARQKFSLLYSIQ	12399
Db	2041	GQYQOMAPKRLARLHYSSSINAMSTKEPFSWIKYDLAPMITHGITQCARQKFSLLYSIQ	21000
Qy	1240	FIIMYSLDGKKMOYIRNRNSTGCTLMVFEFGVNDSSGIKHNIFNPPIIARYIRLHPHYSTRS	12999
Db	2101	FIIMYSLDGKKMOYIRNRNSTGCTLMVFEFGVNDSSGIKHNIFNPPIIARYIRLHPHYSTRS	21600
Qy	1300	TLRRELNGCDLNSCSPMLGEMSKAISDAOITASSYFTNNFATWSPSKARLHLQGRSNAMR	13599
Db	2161	TLRRELNGCDLNSCSPMLGEMSKAISDAOITASSYFTNNFATWSPSKARLHLQGRSNAMR	22200
Qy	1360	PQVNNPKEMLOVDFOKRMKATGYTTOGVKSLLTSMVYKRELLSSQDGHOWTLFPONKV	14199
Db	2221	PQVNNPKEMLOVDFOKRMKATGYTTOGVKSLLTSMVYKRELLSSQDGHOWTLFPONKV	22800
Qy	1420	KVFGNGNDSTFPVYNSLDPLRLRYRIRHPOQSVHQAIALRMEVLGCEAQDLY	1471
Db	2281	KVFGNGNDSTFPVYNSLDPLRLRYRIRHPOQSVHQAIALRMEVLGCEAQDLY	2332
RESULT 10			
US-08-251-937A-4			
: Sequence 4, Application US/08251937A			
: Patent No. 5583209			
GENERAL INFORMATION:			
APPLICANT: Lollar, John S.			
APPLICANT: Runge, Marschall S.			
TITLE OF INVENTION: Hybrid Human/Porcine Factor VIII			
NUMBER OF SEQUENCES: 10			
CORRESPONDENCE ADDRESS:			
ADDRESSEE: Kilpatrick & Cody			
STREET: 1100 Peachtree Street			
CITY: Atlanta			
STATE: Georgia			
COUNTRY: US			
ZIP: 30309			
COMPUTER READABLE FORM:			
MEDIUM TYPE: Floppy disk			
COMPUTER: IBM PC compatible			
OPERATING SYSTEM: PC-DOS/MS-DOS			
SOFTWARE: PatentIn Release #1.0, Version #1.25			
CURRENT APPLICATION DATA:			
APPLICATION NUMBER: US/08/251,937A			
FILING DATE: 31-MAY-1994			
CLASSIFICATION: 435			
PRIOR APPLICATION DATA:			
APPLICATION NUMBER: US 07/864,004			
FILING DATE: 07-APR-1992			
ATTORNEY/AGENT INFORMATION:			
NAME: Pratl, John S.			
REGISTRATION NUMBER: 29,476			



REFERENCE/DOCKET NUMBER: EMU106DIV  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 404-815-6367  
TELEFAX: 404-815-6555  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2332 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHEICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
ORIGINAL SOURCE:  
ORGANISM: Homo sapien  
TISSUE TYPE: Liver cDNA sequence  
US-08-251-937A-4

Query Match 93.0%; Score 7315; DB 1; Length 2332;  
Best Local Similarity 62.3%; Pred. No. 0;  
Matches 1452; Conservative 0; Mismatches 0; Indels 880; Gaps 1;

QY 20 ATRRYLGAVELSMDYMOSDLGELPVDARPPRPVKSPFNTSVYKKTLFVEFTVHLFN 79  
DB 1 ATRRYLGAVELSMDYMOSDLGELPVDARPPRPVKSPFNTSVYKKTLFVEFTVHLFN 60  
QY 80 IAKPRPMGLGPTIOAEVYDVTVVITLKNMASHPVSLHAGVSYWKASEGAEYDDQTSQ 139  
DB 61 IAKPRPMGLGPTIOAEVYDVTVVITLKNMASHPVSLHAGVSYWKASEGAEYDDQTSQ 120  
QY 140 REKEDKAFPGSGHTYVQVLKENGPMASDPLCTLYSLSHVDLYKDLNSGLIGALLVCR 199  
DB 121 REKEDKAFPGSGHTYVQVLKENGPMASDPLCTLYSLSHVDLYKDLNSGLIGALLVCR 180  
QY 200 EGSIAKEVQTLLHKFILFAVFEDEGKSWHSETKNSLMODRDAASARAPMKHTVNGYNR 259  
DB 181 EGSIAKEVQTLLHKFILFAVFEDEGKSWHSETKNSLMODRDAASARAPMKHTVNGYNR 240  
QY 260 SLPLGLGCHRSYVHWIIGMGTTPVHSLFLEGHTFLVRNHRQASLEISPTFLTAQTL 319  
DB 241 SLPLGLGCHRSYVHWIIGMGTTPVHSLFLEGHTFLVRNHRQASLEISPTFLTAQTL 300  
QY 320 MDLQGFLLFCHISSHOHGMAYYVVDSCPEPOLRMKNNEAEYDDDLTDSMDVYRF 379  
DB 301 MDLQGFLLFCHISSHOHGMAYYVVDSCPEPOLRMKNNEAEYDDDLTDSMDVYRF 360  
QY 380 DDDNSPFIQIRSAVAKKHPKTMVHYIAAEEDMDAPLVLPDDRYSKSOYLNNGPORIG 439  
DB 361 DDDNSPFIQIRSAVAKKHPKTMVHYIAAEEDMDAPLVLPDDRYSKSOYLNNGPORIG 420  
QY 440 RKKYKVRMAYTDETFKTRAIQIHESGILGPLLYGEVDTLLIFKNQASRPYNIYPHGI 499  
DB 421 RKKYKVRMAYTDETFKTRAIQIHESGILGPLLYGEVDTLLIFKNQASRPYNIYPHGI 480  
QY 500 TDVPRVLRRLPKGVKHLKDPILPGLGELFKKTYTVADGTPKSPRCLTRYSSFVME 559  
DB 481 TDVPRVLRRLPKGVKHLKDPILPGLGELFKKTYTVADGTPKSPRCLTRYSSFVME 540  
QY 560 RDLASGLGPLLICYSKEVDQGNQIMSDKRNVLISFVDENRSMYLTENIORFLPNPAG 619  
DB 541 RDLASGLGPLLICYSKEVDQGNQIMSDKRNVLISFVDENRSMYLTENIORFLPNPAG 600  
QY 620 VOLEDEPEFOASNIHMSINGYFDSIQLSVCLHEVAAYWILSIGAQDTFLSVFFSGYTERKH 679  
DB 601 VOLEDEPEFOASNIHMSINGYFDSIQLSVCLHEVAAYWILSIGAQDTFLSVFFSGYTERKH 660  
QY 680 KMYVEDTLTLPFSGEYFEMENPGMLIICCHNSDFNRGMTALTKKSSCDKNTGDYFE 739  
DB 661 KMYVEDTLTLPFSGEYFEMENPGMLIICCHNSDFNRGMTALTKKSSCDKNTGDYFE 720  
QY 740 DSYEDISAYLLSKNNAIEPRSFSONSRHPSTROKOFNAT----- 778

DB 721 DSYEDISAYLLSKNNAIEPRSFSONSRHPSTROKOFNATTTIPENDIEKTDPPFAHRTMP 780  
QY 779 ----- 778  
DB 781 KIONVSSDILLMLRQSPTHGLSLSLQEAKEYTFSDPSGAIDSNNSLSEMTFRPQ 840  
QY 779 ----- 778  
DB 841 LHHSGDMVFTPEGLQLRLNEKLGTTAATELKKLDKVSSTSNLISPTPSDNLAAGTDN 900  
QY 779 ----- 778  
DB 901 TSSLGPPMPVHYDSQDLDTTLFGKKSSPLTBSGGLSLSEENNSKLLSGLMNSQESSW 960  
QY 779 ----- 778  
DB 961 GKNVSTESGLFKGKRAHGALLTKDNALFKSISLTKNTKSNSATNKRTHIDGPSL 1020  
QY 779 ----- 778  
DB 1021 LIENSPYWMNLESDEFEKKVPLIHDRMLDKNATLARLHMSNKTTSSKNMEMVOOK 1080  
QY 779 ----- 778  
DB 1081 KEGTLPDQAQNDMSFFKMLPLPESARWIOPTHGKNSLNSGGSPSKOLVSLGPEKSYEG 1140  
QY 779 ----- 778  
DB 1141 QNFLSEKKKVVYKGEFTKDVGLKEMVFPSSRNPLTLNLDNLHENNTNQEKKIOEBIEK 1200  
QY 779 ----- 778  
DB 1201 KETLLQENVVLPQIHTVTGTRKFMKNLFLSTRONVGSYEGAVAPVLODFRSLNDSTNR 1260  
QY 779 ----- 778  
DB 1261 TKKTHAHSKKGEEENLEGLNOTQOIEKYACTTRISPNTSQNFYVQSRKALKQFRL 1320  
QY 779 ----- 778  
DB 1321 PLEETELEKRIIVDDTSTQWSKNMKHLTPSTLTOIDYNEKEGAIQSPSLDCLTRSHSI 1380  
QY 779 ----- 778  
DB 1381 POANRSLPLIAKVSSFSIRPIYTLRYLFDONSHLPASTRKSDGVOESSHFLQAKK 1440  
QY 779 ----- 778  
DB 1441 NNLSLAILTLEMTGQREVSLGTSATNSVYKKYENTVLPKPDLPKTSKGYELLPKYHI 1500  
QY 779 ----- 778  
DB 1501 YOKDLFPTETSGNSPGHLDIVEGSLLOSTEGAIKWNANRPGKVPFLRVATESSAKTPSK 1560  
QY 779 ----- 778  
DB 1561 LLDPLAMDNHYGTQIPKEEMKSQEKSPKTAFAKKKOTLLSNACESNHAIAAINEGQNK 1620  
QY 779 ----- 778  
DB 1621 ELEVYMAOGTRELCSQNPVYLKRHOREITRTTIQSQOEELDDDTISVEKKKEDPDIY 1680  
QY 820 DEDENQSPRSFOKKTIRHYFIAAVERLMDYGMSSSPHYLRNRAQSGSVPOFKKVVFOEFTD 879  
DB 1681 DEDENQSPRSFOKKTIRHYFIAAVERLMDYGMSSSPHYLRNRAQSGSVPOFKKVVFOEFTD 1740  
QY 880 GSFQPLVRCGLNEHLGLGPIYIRAEVEDNIMVPRNARSPPYSGLISYEEDOROGA 939  
DB 1741 GSFQPLVRCGLNEHLGLGPIYIRAEVEDNIMVPRNARSPPYSGLISYEEDOROGA 1800  
QY 940 EPRKNFVKNETKTYFWKVVQHMAPTKDEFDCKAMAYFSVDLLEKDVHSGILGPLVCHT 999



```

Db 1801 EBRKAFVKNETKTYFWKVQHHMAPTKDEFDCKAMAYFSDVDLEKDVHSLGIPLLVCHT 1860
QY 1000 NTLNPAHGRQVYVOEBALEFTTFDETKSWYFTENMERNCAPNIOEMDETFEKENYRFA 1059
Db 1861 NTLNPAHGRQVYVOEBALEFTTFDETKSWYFTENMERNCAPNIOEMDETFEKENYRFA 1920
QY 1060 INGYIMDTLPGLVMAODORIRMYLLSMGSNENIHSHFSGHVFTVKKKEEYKMALYNLYP 1119
Db 1921 INGYIMDTLPGLVMAODORIRMYLLSMGSNENIHSHFSGHVFTVKKKEEYKMALYNLYP 1980
QY 1120 GVEFEVEMPSKAGIMRVCLIGEHLLHAGMSTFLVYSNKCQPLCMASHINDPQITAS 1179
Db 1981 GVEFEVEMPSKAGIMRVCLIGEHLLHAGMSTFLVYSNKCQPLCMASHINDPQITAS 2040
QY 1180 GQYGWAPKLARLHYSGSINAMSTKEPFSWIKVDLAPMIIHIGTKOAROKFSSYISQ 1239
Db 2041 GQYGWAPKLARLHYSGSINAMSTKEPFSWIKVDLAPMIIHIGTKOAROKFSSYISQ 2100
QY 1240 FIMYSLDGKKWQTYRGNSTGTLLMVEFGVNDSSGIKHNIFNPPIIARYIRLHPTHSIRS 1299
Db 2101 FIMYSLDGKKWQTYRGNSTGTLLMVEFGVNDSSGIKHNIFNPPIIARYIRLHPTHSIRS 2160
QY 1300 TLRMLMGCDLNSGSMPLGMEKSAISDAQITASSYFTNMFATWSPSKARLHLOGRSNAMR 1359
Db 2161 TLRMLMGCDLNSGSMPLGMEKSAISDAQITASSYFTNMFATWSPSKARLHLOGRSNAMR 2220
QY 1360 PQVNNPKEMLOVDFOKTMKVGTGTTQGVKSLTSMYKEFLISSODGQHWTLFPONGKY 1419
Db 2221 PQVNNPKEMLOVDFOKTMKVGTGTTQGVKSLTSMYKEFLISSODGQHWTLFPONGKY 2280
QY 1420 KVFQGNOSFTFVNSLDPPLLTRILRIHPQSVHQAIALRMEVLGCEADLY 1471
Db 2281 KVFQGNOSFTFVNSLDPPLLTRILRIHPQSVHQAIALRMEVLGCEADLY 2332

RESULT 11
US-08-212-133A-2
: Sequence 2, Application US/08212133A
: Patent No. 5663060
: GENERAL INFORMATION:
: APPLICANT: Lollar, John S.
: APPLICANT: Runge, Marshall S.
: TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
: NUMBER OF SEQUENCES: 12
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Kilpatrick & Cody
: STREET: 100 Peachtree Street
: City: Atlanta
: STATE: Georgia
: COUNTRY: US
: ZIP: 30303
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: IBM PC compatible
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/212.133A
: FILING DATE: March 11, 1994
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/864,004
: FILING DATE: 07-APR-1992
: ATTORNEY/AGENT INFORMATION:
: NAME: Pabst, Patrea L.
: REGISTRATION NUMBER: 31,284
: REFERENCE/DOCKET NUMBER: EMU/76677
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 404-572-6555
: TELEFAX: 404-572-6555
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2332 amino acids

```

```

: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: HYPOTHEICAL: YES
: ANTI-SENSE: NO
: FRAGMENT TYPE: N-terminal
: ORIGINAL SOURCE:
: ORGANISM: Homo sapien
: TISSUE TYPE: Liver cDNA sequence
: US-08-212-133A-2

```

```

Query Match          93.0%; Score 7315; DB 1; Length 2332;
Best Local Similarity 62.3%; Pred. No. 0;
Matches 1452; Conservative 0; Mismatches 0; Indels 880; Gaps 1;

```

```

QY 20 ATRRRYLGAVELSDWYMQSDCLBLYDARFPPRVKSPFPNTSVYKKTLVEEFTVHLFN 79
Db 1 ATRRRYLGAVELSDWYMQSDCLBLYDARFPPRVKSPFPNTSVYKKTLVEEFTVHLFN 60
QY 80 IAKRPMMGLIGPTQAEVYPTVTTLKNMASHVSLHAGVSYWKASGAEYDQTSQ 139
Db 61 IAKRPMMGLIGPTQAEVYPTVTTLKNMASHVSLHAGVSYWKASGAEYDQTSQ 120
QY 140 REKEDKVEPPGSHYVWQVLKENGPMASDPLCLTYSTLSHVDLVKDNSGLIGALVCR 199
Db 121 REKEDKVEPPGSHYVWQVLKENGPMASDPLCLTYSTLSHVDLVKDNSGLIGALVCR 180
QY 200 EGSIAKEKQTLHKFTLLFAVDEGKSWHSERKNSLMODROAASRAMPKMTVNGVYNR 259
Db 181 EGSIAKEKQTLHKFTLLFAVDEGKSWHSERKNSLMODROAASRAMPKMTVNGVYNR 240
QY 260 SLPGILGCHRRKSYVHVIGMGTPEVHSIFLEGHTFLVFNHQAASLESPITFLAQTLL 319
Db 241 SLPGILGCHRRKSYVHVIGMGTPEVHSIFLEGHTFLVFNHQAASLESPITFLAQTLL 300
QY 320 MDIGQFLFCCHISSHQHDMGEAYVYKVDSCPEEPQLRMKNNBEAEDYDDDLTDSMDVYRF 379
Db 301 MDIGQFLFCCHISSHQHDMGEAYVYKVDSCPEEPQLRMKNNBEAEDYDDDLTDSMDVYRF 360
QY 380 DDDNSPSFIQISVAKKHRTVHTYIAAEEEDMDYAPLYLADDDNSYSQYILNKPQIRG 439
Db 361 DDDNSPSFIQISVAKKHRTVHTYIAAEEEDMDYAPLYLADDDNSYSQYILNKPQIRG 420
QY 440 RKTKKVRMAVYDEPFTKEATQIHESGILGLYGEVGTLLITRKNASRPYNTYPRGI 499
Db 421 RKTKKVRMAVYDEPFTKEATQIHESGILGLYGEVGTLLITRKNASRPYNTYPRGI 480
QY 500 TDVRPLYSRRLPKGVKHLDPILPGLIFKYKWTYVEDGPTKSDPRLTRYSSFVME 559
Db 481 TDVRPLYSRRLPKGVKHLDPILPGLIFKYKWTYVEDGPTKSDPRLTRYSSFVME 540
QY 560 RDLASGLIPPLICYKESVDQGNQIMSDKRVYLLFSVDEKRSYLTENIQRFIPNAG 619
Db 541 RDLASGLIPPLICYKESVDQGNQIMSDKRVYLLFSVDEKRSYLTENIQRFIPNAG 600
QY 620 VOLEDEFOASINIMHSINGVYFDSIQLSVCLHEVAWYILSIGAOTDLSVFFSGYTRKH 679
Db 601 VOLEDEFOASINIMHSINGVYFDSIQLSVCLHEVAWYILSIGAOTDLSVFFSGYTRKH 660
QY 680 KMYVEDTLTLFPFSGEYVFMENPGLWILGCHNSDFRNRGTALLKVSCKDKNTGYYE 739
Db 661 KMYVEDTLTLFPFSGEYVFMENPGLWILGCHNSDFRNRGTALLKVSCKDKNTGYYE 720
QY 740 DSYEDISATLLSKNNAIERFSQNSRHPSTRQKQPNAT----- 778
Db 721 DSYEDISATLLSKNNAIERFSQNSRHPSTRQKQPNAT----- 780
QY 779 ----- 778
Db 781 KIQNVSSSDLLMLRKOSPPTPHGLSLSDQEAKEYTFSDPSBGAIDSNNSLSEMTHTRRPQ 840

```



```

Qy 779 ----- 778
Db 841 LHSQDMVETPESGLQLRLNEKLGTTAATBLKLDKFKVSVTSNNLSTIPSDNLAGTDN 900
Qy 779 ----- 778
Db 901 TSSLCPPSPVHYDSQDLDTTLFGKKSSPLTSGGPLSSEENNDKLESGLMNSQESSW 960
Qy 779 ----- 778
Db 961 GKNVSTESGRLFGKRAHGAPALLTKDNALFKVSTISLKTNTSNNSATNKRKTHIDGPSL 1020
Qy 779 ----- 778
Db 1021 LIENSPSWQNLIESDPFEKKVTPLIHDRMLMDKNATALRLNHSNKTTSSKNEMVQOK 1080
Qy 779 ----- 778
Db 1081 KEGPIPPDAQNDMSFFKMLFLPESARWIOETHGKNSLNSGQSPKQVSLGPEKSVEG 1140
Qy 779 ----- 778
Db 1141 QNPLSEKNVVGGEFTKDVGLKEMVPPSSRNLELTNLDNLHENNTHNOEKKIOEIEK 1200
Qy 779 ----- 778
Db 1201 KETLIGENVVLPQIHVYTGTKNEMKNLFLSTRONVSGYEGATAPVLQDFRSLNDSTNR 1260
Qy 779 ----- 778
Db 1261 TKKHTAHSKKEEENLEGLGNTQKQIVEKYACTTRISPTNSQONFYQRSKRALKQFRL 1320
Qy 779 ----- 778
Db 1321 PLEETELEKRIIVDTSTQMSKNNKHLTPSTLQIDYNEKEKAIQSPSLDCLTRSHSI 1380
Qy 779 ----- 778
Db 1381 PQANRSPPLIAKVSSPFSIRIYILTRVLPQDNSSHLPASVYRKKDSGVQESSHFLQAKK 1440
Qy 779 ----- 778
Db 1441 NNLSLAILTEMTGDQREVSGISATNSVYTKKVENTVLPKPDLPKTSKVELLPKVIH 1500
Qy 779 ----- 778
Db 1501 YOKDLPTETNSGSPGHLDIYEGSLQGTGGAIKMNEANRPKVPFLRVATESSAKTPSK 1560
Qy 779 ----- 778
Db 1561 LLDPLAMDNHGTQIPKEEMKSQEKSPKTAFFKKDTILSNACESNHAIAINEGONKP 1620
Qy 779 ----- 819
Db 1621 EIEYTWAKQRTERLCSQNPVPLKRHOREITRTLQSDQEBIDYDDTISVEMKKEDFDIY 1680
Qy 820 DEDNOSPRROKTRHFIYAVERIMDYGMSSPHVLRNNAOGSGVQFKKVVPOETD 879
Db 1681 DEDEQOSPRROKTRHFIYAVERIMDYGMSSPHVLRNNAOGSGVQFKKVVPOETD 1740
Qy 880 GSFQPLYRGELNHLGLGPIYIRAEVEDNIMVTFERNQASRPYSFYSSLIYEEDQROGA 939
Db 1741 GSFQPLYRGELNHLGLGPIYIRAEVEDNIMVTFERNQASRPYSFYSSLIYEEDQROGA 1800
Qy 940 EPRNNEFKVNETKTYFKVQVHHMAPTKDEFCCKAMAFSDVDLEKDVHSGLIGPLVCHT 999
Db 1801 EPRNNEFKVNETKTYFKVQVHHMAPTKDEFCCKAMAFSDVDLEKDVHSGLIGPLVCHT 1860
Qy 1000 NTLNPAHGRQVTOEFALEFIIPEETKSWYFTEENNERCRAPCNINOMEDPFKEKNYRPHA 1059
Db 1861 NTLNPAHGRQVTOEFALEFIIPEETKSWYFTEENNERCRAPCNINOMEDPFKEKNYRPHA 1920
Qy 1060 INGYIMDTLPGLVMAODORIRWYLLSMGSENINSHSFGHVFTRKKEEYKMALYNLYP 1119

```

```

Db 1921 INGYIMDTLPGLVMAODORIRWYLLSMGSENINSHSFGHVFTRKKEEYKMALYNLYP 1980
Qy 1120 GVFETVEMLPBSKAGIWRRECLIGEHLHAGMSTFLYVSNCKOTPLGMAAGIIRFOITAS 1179
Db 1981 GVFEVEMLPBSKAGIWRRECLIGEHLHAGMSTFLYVSNCKOTPLGMAAGIIRFOITAS 2040
Qy 1180 GYOGOMAPKLARLHYSGSINMSTKEPFSWIKVDLAPMIITHGIKTOGAROKFSSLYTSQ 1239
Db 2041 GYOGOMAPKLARLHYSGSINMSTKEPFSWIKVDLAPMIITHGIKTOGAROKFSSLYTSQ 2100
Qy 1240 FIIMYSLDGKKWQYRGNGSTGLMWFEGNVDSGKIHAFNPPIIARIIRLPHYSIRS 1299
Db 2101 FIIMYSLDGKKWQYRGNGSTGLMWFEGNVDSGKIHAFNPPIIARIIRLPHYSIRS 2160
Qy 1300 TLRMELMGCDLNSCSPMPGMSKAISSAQITASSYFTNMFATWSPSKARLHLOGRSNMR 1359
Db 2161 TLRMELMGCDLNSCSPMPGMSKAISSAQITASSYFTNMFATWSPSKARLHLOGRSNMR 2220
Qy 1360 POWNPKEMLOVDPOKTKKVTGVTQGVKSILTSYVKEFLLSSQDGHQMTLFFQNGKV 1419
Db 2221 POWNPKEMLOVDPOKTKKVTGVTQGVKSILTSYVKEFLLSSQDGHQMTLFFQNGKV 2280
Qy 1420 KVFQGNDSFTPPVNSLDPPLTRYLRIHPOSWVHQAIRMEVLGCEAODLY 1471
Db 2281 KVFQGNDSFTPPVNSLDPPLTRYLRIHPOSWVHQAIRMEVLGCEAODLY 2332

```

```

RESULT 12
US-08-474-503-2
: Sequence 2, Application US/08474503
: Patent No. 5744446
:
: GENERAL INFORMATION:
: APPLICANT: Emory University
: TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
: NUMBER OF SEQUENCES: 12
: CORRESPONDENCE ADDRESSES:
: ADDRESSEE: Kilpatrick & Cody
: STREET: 1100 Peachtree Street, Suite 2800
: CITY: Atlanta
: STATE: Georgia
: COUNTRY: US
: ZIP: 30309
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent in Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/474,503
: FILING DATE: 07-JUN-1995
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Pratt, John S.
: REGISTRATION NUMBER: 29,476
: REFERENCE/DOCKET NUMBER: EMU106CIP(3)
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 404-815-6500
: TELEFAX: 404-815-6555
:
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2332 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: HYPOTHEetical: YES
: ANTI-SENSE: NO
: FRAGMENT TYPE: N-terminal
: ORIGINAL SOURCE:
: ORGANISM: Homo sapien
: TISSUE TYPE: Liver cDNA sequence
:
: US-08-474-503-2

```







QY 1240 FIIMYSLDGKKQWYRGNSGTGLWVFGNVDSGIGKINIFNPPIIARYIRLPHYSIRS 1239  
 |||||||  
 Db 2101 FIIMYSLDGKKQWYRGNSGTGLWVFGNVDSGIGKINIFNPPIIARYIRLPHYSIRS 2160  
 |||||||  
 QY 1300 TLRMELMCGDLNSCMLPGMESKASIDAOITASSYFTNMFATWSPSKARLHLQGRSNMR 1359  
 |||||||  
 Db 2161 TLRMELMCGDLNSCMLPGMESKASIDAOITASSYFTNMFATWSPSKARLHLQGRSNMR 2220  
 |||||||  
 QY 1360 POWNPKEMLOVDPOKTKMKVGVYTGQYKSLTSMYVKEFLISSQDQHQMTLFFQNGKV 1419  
 |||||||  
 Db 2221 POWNPKEMLOVDPOKTKMKVGVYTGQYKSLTSMYVKEFLISSQDQHQMTLFFQNGKV 2280  
 |||||||  
 QY 1420 KVFQGNQDSFTPVVNSLDPLFTRLRIRHPOSWVHQLRMREVLCGEAODLY 1471  
 |||||||  
 Db 2281 KVFQGNQDSFTPVVNSLDPLFTRLRIRHPOSWVHQLRMREVLCGEAODLY 2332  
 |||||||  
 RESULT 13  
 US-08-670-707A-2  
 : Sequence 2, Application US/08670707A  
 : Patent No. 5859204  
 : GENERAL INFORMATION:  
 : APPLICANT: Lohlar, John S.  
 : TITLE OF INVENTION: Hybrid Human/Animal Factor VIII  
 : NUMBER OF SEQUENCES: 40  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: Greenlee, Winnet and Sullivan, P.C.  
 : STREET: 5370 Manhattan Circle Suite 201  
 : CITY: Boulder  
 : STATE: Colorado  
 : COUNTRY: USA  
 : ZIP: 80303  
 : COMPUTER READABLE FORM:  
 : MEDIUM TYPE: floppy disk  
 : OPERATING SYSTEM: PC-DOS/MS-DOS  
 : SOFTWARE: Patent Release #1.0, Version #1.30  
 : CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: US/08/670,707A  
 : FILING DATE: 26-JUN-1996  
 : CLASSIFICATION: 435  
 : PRIOR APPLICATION DATA:  
 : APPLICATION NUMBER: WO PCT/US94/13200  
 : FILING DATE: 15-NOV-1994  
 : PRIOR APPLICATION DATA:  
 : APPLICATION NUMBER: US 08/212,133  
 : FILING DATE: 11-MAR-1994  
 : PRIOR APPLICATION DATA:  
 : APPLICATION NUMBER: US 07/864,004  
 : FILING DATE: 07-APR-1992  
 : ATTORNEY/AGENT INFORMATION:  
 : NAME: Greenlee, Lorance L.  
 : REGISTRATION NUMBER: 27,894  
 : REFERENCE/DOCKET NUMBER: 75-95F  
 : TELECOMMUNICATION INFORMATION:  
 : TELEPHONE: 303/499-8080  
 : TELEFAX: 303/499-8089  
 : INFORMATION FOR SEQ ID NO: 2:  
 : SEQUENCE CHARACTERISTICS:  
 : LENGTH: 2332 amino acids  
 : TYPE: amino acid  
 : STRANDEDNESS: single  
 : TOPOLOGY: not relevant  
 : MOLECULE TYPE: protein  
 : HYPOTHEICAL: YES  
 : ANTI-SENSE: NO  
 : FRAGMENT TYPE: N-terminal  
 : ORIGINAL SOURCE:  
 : ORGANISM: Homo sapiens  
 : TISSUE TYPE: Liver  
 : US-08-670-707A-2

Query Match 93.0%; Score 7315; DB 2; Length 2332;  
 Best Local Similarity 62.3%; Pred. No. 0;  
 Matches 1452; Conservative 0; Mismatches 0; Indels 880; Gaps 1;  
 QY 20 ATRRYYGAVELSDMYDMSDGLGELPVDARPPRPVRSFPFTSVYKKTLEVEFTVHLFN 79  
 |||||||  
 Db 1 ATRRYYGAVELSDMYDMSDGLGELPVDARPPRPVRSFPFTSVYKKTLEVEFTVHLFN 60  
 |||||||  
 QY 80 IAKRPPRMGLGPTIOAEVYDVTYITLKNNASHPVSLAHAVGYWKASBGAEYDDOTSQ 139  
 |||||||  
 Db 61 IAKRPPRMGLGPTIOAEVYDVTYITLKNNASHPVSLAHAVGYWKASBGAEYDDOTSQ 120  
 |||||||  
 QY 140 REKDDKVPFGSGHTYVQVLKENGPMASDPLCLTYSLSHVDYVKDLSGLTALLVCR 199  
 |||||||  
 Db 121 REKDDKVPFGSGHTYVQVLKENGPMASDPLCLTYSLSHVDYVKDLSGLTALLVCR 180  
 |||||||  
 QY 200 EGSIAKERTQTLHKFILLFAVEDEGKSWHSETKNSLMODRDAASARAMPKHTYNGVYNR 259  
 |||||||  
 Db 181 EGSIAKERTQTLHKFILLFAVEDEGKSWHSETKNSLMODRDAASARAMPKHTYNGVYNR 240  
 |||||||  
 QY 260 SLPLGIGCHRSYVWHYIGMGTPEVHSIFLEGHTFLVNRHQASLEISPTFLTAOTLL 319  
 |||||||  
 Db 241 SLPLGIGCHRSYVWHYIGMGTPEVHSIFLEGHTFLVNRHQASLEISPTFLTAOTLL 300  
 |||||||  
 QY 320 MDLGOFLLFCHISSHOHDMGEMAYVVDSCPEEPLRMKNNEADYDDDLTDSMDVYRF 379  
 |||||||  
 Db 301 MDLGOFLLFCHISSHOHDMGEMAYVVDSCPEEPLRMKNNEADYDDDLTDSMDVYRF 360  
 |||||||  
 QY 380 DDDNSPSFIQIRYAKKHPKTPWVHYIAAEEEDMYPAPLYLAPDDRYSKQYLLNNGPORIG 439  
 |||||||  
 Db 361 DDDNSPSFIQIRYAKKHPKTPWVHYIAAEEEDMYPAPLYLAPDDRYSKQYLLNNGPORIG 420  
 |||||||  
 QY 440 RKYKVRPMATYDTEFTTREAIQHESGILGPLLYGEVDTLLITFKQASRPYNIYPHGI 499  
 |||||||  
 Db 421 RKYKVRPMATYDTEFTTREAIQHESGILGPLLYGEVDTLLITFKQASRPYNIYPHGI 480  
 |||||||  
 QY 500 TDVAPLVSRLPKGVKHLKDPPLIGELFKYKWTYVEDGTSKDPCLTFYSSPFNME 559  
 |||||||  
 Db 481 TDVAPLVSRLPKGVKHLKDPPLIGELFKYKWTYVEDGTSKDPCLTFYSSPFNME 540  
 |||||||  
 QY 560 RDLASGLIGPLLYCKESVDQGNQIMSDKRNVLFSVFENRSWYLTENIQRELPNPAQ 619  
 |||||||  
 Db 541 RDLASGLIGPLLYCKESVDQGNQIMSDKRNVLFSVFENRSWYLTENIQRELPNPAQ 600  
 |||||||  
 QY 620 VQLEDPERQASNMHSGINGYFDSLQSVCLHEVAWYIISGAQDPLSFVSGGYTFKH 679  
 |||||||  
 Db 601 VQLEDPERQASNMHSGINGYFDSLQSVCLHEVAWYIISGAQDPLSFVSGGYTFKH 660  
 |||||||  
 QY 680 KMVEEDTLTLPFSGETVPFMSMENPGMLIGCHNSDRNKGMTALLKVSQDKMTGDTYE 739  
 |||||||  
 Db 661 KMVEEDTLTLPFSGETVPFMSMENPGMLIGCHNSDRNKGMTALLKVSQDKMTGDTYE 720  
 |||||||  
 QY 740 DSYEDISAYLLSKNNAIEPRFSQNSRHPSTRQKQFAT 778  
 |||||||  
 Db 721 DSYEDISAYLLSKNNAIEPRFSQNSRHPSTRQKQFAT 780  
 |||||||  
 QY 779 778  
 |||||||  
 Db 781 KIQNVSSDILLMLROSPTPHGLSLDLOEAKYETFSDDSPGAIIDSNNLSIEMTHRPQ 840  
 |||||||  
 QY 779 778  
 |||||||  
 Db 841 LHMSGDMVTPPESGLOLRLEKLGTTATATLKLDFVVSSTNNLISTIPSDNLAAGTDN 900  
 |||||||  
 QY 779 778  
 |||||||  
 Db 901 TTSILGPPSMVHYHDSQDITLLFKGKSSPLTESGGLSLSEENNSKLLSEGLMNSQSSW 960  
 |||||||  
 QY 779 778  
 |||||||  
 Db 961 GKNVSTESGRLFKGRAHGPALLTKDNALFKVISILKTKNTKTSNNSATNRKTHIDPSL 1020  
 |||||||  
 QY 779 778



Db 1021 LIENSPVWONILESDETEKVVPLIHDMMLDKNATALRLNMSNKTSSKNMEVQOK 1080  
QY 779 ----- 778  
Db 1081 KEGPIPDQONPDMSEFFKMLFIPESARMIQRTGKNSLNSGOGPSPKQLVSLGPEKSVEG 1140  
QY 779 ----- 778  
Db 1141 QNPLSEKKNVVYKGEFTKDVCKEMVFPSSRNFLTLNLDNLHENNTNHQEKIOEIEIK 1200  
QY 779 ----- 778  
Db 1201 KETLIQENVVLPQIHVTYGTKNFMKNLFLSTRONVEGSEGAAPVLDQFRSLNDSTNR 1260  
QY 779 ----- 778  
Db 1261 TKKHTAHFSKSGEENLEGLNOTQOIVEKACTTRISPNTSQONFVYORSKRAIKQFRL 1320  
QY 779 ----- 778  
Db 1321 PLEETELEKRIIVDDTSTOWSKMKHLTPSTLTQIDYNEKEGATQSPUSDCLFRSHSI 1380  
QY 779 ----- 778  
Db 1381 PQANRSPPLIAKVSSFPSIRPIYLRVLFDQNSSHLPAASYRKKDQVQESSHFLQGAKK 1440  
QY 779 ----- 778  
Db 1441 NNLSLAILTEMTGDQREVSGISGTSATNSVYTKKVENTVLPKPDLPKTSKVELLPKVIH 1500  
QY 779 ----- 778  
Db 1501 YQKDLFPTETSGSPGHLDIVEGSLLOGTGCAIKWNEANRPGKVPFLRVATESAKTPSK 1560  
QY 779 ----- 778  
Db 1561 LLDPLAMDNYGTQIPKEEMKSQEKSPKTAKKKDTILSLNACESNHAIAAINEGQMKP 1620  
QY 779 ----- 819  
Db 1621 EIEYTWAKOGTERLCSQNPFLKRNHOREITPTLOSQDEEIDYDPTISVEKKKEDFDY 1680  
QY 820 DEDENQSPRSFOKTRAHYFIAAVERLMDYGMSSSPHYLNNRAQSGSVPOFKKVFQOEFTD 879  
Db 1681 DEDENQSPRSFOKTRAHYFIAAVERLMDYGMSSSPHYLNNRAQSGSVPOFKKVFQOEFTD 1740  
QY 880 GSFTQPLRYGELNHLGLGPIYIRAEVEDNTMVFTRNOASRPYSYSSLSIEEDQROGA 939  
Db 1741 GSFTQPLRYGELNHLGLGPIYIRAEVEDNTMVFTRNOASRPYSYSSLSIEEDQROGA 1800  
QY 940 EPRKNFYKPNTEKTYFMKVQHMMAPTKDEDFCKAWAYESDVLDLEKDVHSGILGPLVCHT 999  
Db 1801 EPRKNFYKPNTEKTYFMKVQHMMAPTKDEDFCKAWAYESDVLDLEKDVHSGILGPLVCHT 1860  
QY 1000 NTLNPAHGRQVTVQDFALFETTFDETKSMYFETENNERNCRAPCNQIOMEDPTFKENYRFA 1059  
Db 1861 NTLNPAHGRQVTVQDFALFETTFDETKSMYFETENNERNCRAPCNQIOMEDPTFKENYRFA 1920  
QY 1060 INGYIMDTLPGLVMAODORIRMYLLSMGSENENHSHFSGHVFYRKKEEYKMAALYNLXP 1119  
Db 1921 INGYIMDTLPGLVMAODORIRMYLLSMGSENENHSHFSGHVFYRKKEEYKMAALYNLXP 1980  
QY 1120 GVEFTVEMLPCKAGIWRRECLIGEHLHAGMSTFLVYSNKCOTPLGMAAGHTRDQITAS 1179  
Db 1981 GVEFTVEMLPCKAGIWRRECLIGEHLHAGMSTFLVYSNKCOTPLGMAAGHTRDQITAS 2040  
QY 1180 GQYGQMAPKLARLHYSGSINAMSTKEPFSWIKVDLLAPMIHIGIKTOGAROKFSSLYISQ 1239  
Db 2041 GQYGQMAPKLARLHYSGSINAMSTKEPFSWIKVDLLAPMIHIGIKTOGAROKFSSLYISQ 2100  
QY 1240 FIIMYSLDGKKWQTYRGNSGTGLMVFEGVNDSSGIKHNIFNPPIIARIYIRLHPHTYSIRS 1299

Db 2101 FIIMYSLDGKKWQTYRGNSGTGLMVFEGVNDSSGIKHNIFNPPIIARIYIRLHPHTYSIRS 2160  
QY 1300 TLRMELMGCDLNSCMPLGMEKSAISDAQITPASSYFTNMFAFTWSPSKARLHLOGRSNMR 1359  
Db 2161 TLRMELMGCDLNSCMPLGMEKSAISDAQITPASSYFTNMFAFTWSPSKARLHLOGRSNMR 2220  
QY 1360 POVNNPKEMVLQVDFQKTKMYGVTTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKV 1419  
Db 2221 POVNNPKEMVLQVDFQKTKMYGVTTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKV 2280  
QY 1420 KVFQGNQDSFTVNVNSLDPPILTRYLRIHQPSWVHQIALRMEVLGCEKADLY 1471  
Db 2281 KVFQGNQDSFTVNVNSLDPPILTRYLRIHQPSWVHQIALRMEVLGCEKADLY 2332

RESULT 14  
US-09-037-601-2  
; Sequence 2, Application US/09037601  
; Patent No. 6180371  
; GENERAL INFORMATION:  
; APPLICANT: Lollar, John S.  
; TITLE OF INVENTION: Hybrid Human/Animal Factor VIII  
; NUMBER OF SEQUENCES: 40  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.  
; STREET: 5370 Manhattan Circle Suite 201  
; CITY: Boulder  
; STATE: Colorado  
; COUNTRY: USA  
; ZIP: 80303  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/037,601  
; FILING DATE: 26-JUN-1996  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: WO PCT/US94/13200  
; FILING DATE: 15-NOV-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/212,133  
; FILING DATE: 11-MAR-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/864,004  
; FILING DATE: 07-APR-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ferber, Donna M.  
; REGISTRATION NUMBER: 33,878  
; REFERENCE/DOCKET NUMBER: 75-95F  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 303/499-8080  
; TELEFAX: 303/499-8089  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2332 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: not relevant  
; MOLECULE TYPE: protein  
; HYDROTHERICAL: YES  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: N-terminal  
; ORIGINAL SOURCE:  
; ORGANISM: Homo sapiens  
; TISSUE TYPE: Liver  
; US-09-037-601-2

Query Match 93.0%; Score 7315; DB 4; Length 2332;  
Best Local Similarity 62.3%; Pred. No. 0;



Matches 1452; Conservative 0; Mismatches 0; Indels 880; Gaps 1;

```
OY 20 ATRRYLGAVELSMYDMOSDGLGELPVDARFPBRVPKSPFENSVYKKTLFVEFVHLFN 79
Db 1 ATRRYLGAVELSMYDMOSDGLGELPVDARFPBRVPKSPFENSVYKKTLFVEFVHLFN 60
OY 80 IAKPRPMWGLGPTIOAEVYDTVVITLKNMASHPVSLHAGVSVYMKASEGAEYDQTSQ 139
Db 61 IAKPRPMWGLGPTIOAEVYDTVVITLKNMASHPVSLHAGVSVYMKASEGAEYDQTSQ 120
OY 140 REKEDKVPFGSGSHYVWQVLKENGPMASDPLCLTYSLSHVDLVKDLNSGLIGALLVCR 199
Db 121 REKEDKVPFGSGSHYVWQVLKENGPMASDPLCLTYSLSHVDLVKDLNSGLIGALLVCR 180
OY 200 EBSLAKKETOTLHKFTLLFAVFEDEKSWHSETKNLMDROAASARAPKMTHTVGYNR 259
Db 181 EBSLAKKETOTLHKFTLLFAVFEDEKSWHSETKNLMDROAASARAPKMTHTVGYNR 240
OY 260 SLPLGLGCHRSVYWHVIGMGTTPREVHSIFLEGHTFLVRNHRQASLEISPTIFLTAQTLL 319
Db 241 SLPLGLGCHRSVYWHVIGMGTTPREVHSIFLEGHTFLVRNHRQASLEISPTIFLTAQTLL 300
OY 320 MDLGFLLFCHISSHQHDGMEAYYKVDSCPEBPQLRMKNNEEAEDYDDLTDSEMDVYRF 379
Db 301 MDLGFLLFCHISSHQHDGMEAYYKVDSCPEBPQLRMKNNEEAEDYDDLTDSEMDVYRF 360
OY 360 DDDNSPSTIOTRSVAKKHKPKMTWVHTAAEEBMDVAPLYLAPDDSYKSQYLNNPQRIG 439
Db 361 DDDNSPSTIOTRSVAKKHKPKMTWVHTAAEEBMDVAPLYLAPDDSYKSQYLNNPQRIG 420
OY 440 RYKVKVRPMATYDEFKTRREALQHESGILGPLLYGEVDTLLIFKNOASRPYNIYPHGI 499
Db 421 RYKVKVRPMATYDEFKTRREALQHESGILGPLLYGEVDTLLIFKNOASRPYNIYPHGI 480
OY 500 TDVRLPLYSRLPKGVKHLKDPILLGEITFKYKWTVYEDGPTKSDPRLTRYSSFVME 559
Db 481 TDVRLPLYSRLPKGVKHLKDPILLGEITFKYKWTVYEDGPTKSDPRLTRYSSFVME 540
OY 560 RDLASGLGPLLICYSKVSDRGNOIMSDKRNVLISVFEDENRSYLTENIQRLPNPAG 619
Db 541 RDLASGLGPLLICYSKVSDRGNOIMSDKRNVLISVFEDENRSYLTENIQRLPNPAG 600
OY 620 VOLDEPFEQASINMHSINGVYFDSLQLSVCLHEVAYWYILSIGAOTDPLSFFSGYTFKH 679
Db 601 VOLDEPFEQASINMHSINGVYFDSLQLSVCLHEVAYWYILSIGAOTDPLSFFSGYTFKH 660
OY 660 KMYVEDTLTLPFSGEIVFMSMENPGLIILGCHNSDFRNRGNTALLKYSSCDKNTGDYFE 739
Db 661 KMYVEDTLTLPFSGEIVFMSMENPGLIILGCHNSDFRNRGNTALLKYSSCDKNTGDYFE 720
OY 740 DSYEDISAYLISKNNALIPRFSFSONSRHPSTROKOFNAT 778
Db 721 DSYEDISAYLISKNNALIPRFSFSONSRHPSTROKOFNAT 780
OY 779 778
Db 781 KIQWSSSSDMLMLRQSTPHGLSLSDQLEAKYETFSDDPSGATDSNNSLSEMTHTFRPQ 840
OY 779 778
Db 841 LHSQDMVFPESGLOLRNLEKLTGTAATELKLDFKVSSTNNLSTIPSDNLAGTDN 900
OY 779 778
Db 901 TSSLGPPSMVHYDSQDLDTLFGKKSSPLTESGGLSSEENNSKILLESGLMNSQESSW 960
OY 779 778
Db 961 GKNVSTESGRLFKGRAGHAPALLTKDNALFKVSTISLKTAKTSNNSATNKRTHLDGPSL 1020
OY 779 778
Db 1021 LIENSPVWONLIESDTEFFKVTPLIHDRMLMDKNATALRLHMSNKTTSKNNMEVQOK 1080
```

```
OY 779 778
Db 1081 KEGTLPDAQNDMSFEKMLPLPESARMIQRTHGKNSLNSGOGSPKQVSLGPKBSVEG 1140
OY 779 778
Db 1141 QNFLSEKKNVYKGEFTKDVGLKEMVFPSSRNFLTNLNDLHENNTNNOEKKIOEELIEK 1200
OY 779 778
Db 1201 KETLQENVVLPQIHTVYTRKNFMKLEFLSTRQWEGSYEGAYAPVQDERSLNDSTNR 1260
OY 779 778
Db 1261 TKKHTAHSKKGEBELEGLOTKQIYEKYACTTRISPNTSQONFVYORSKRALKQERL 1320
OY 779 778
Db 1321 PLEETELEKRIIVDDTSTQWKSNNKHLPLPSTLQIDYNEKEKGATTSQPLSDCLTRSHSI 1380
OY 779 778
Db 1381 POANRSPPLIAKVSFPIRPIYLTRVLFDONSSHLPAAZYRKDKDSGVQESSHFLQAKK 1440
OY 779 778
Db 1441 NNLSLAILTLEKTDQREVSLGTSATNSVYTKKVENTVLPKPDLPKTSKGYELLPKVHI 1500
OY 779 778
Db 1501 YQKDLFPETNSGSPGHLDIVBGSLLQGTGEGAIKKNENNRQKVPFLVATESSAKTBSK 1560
OY 779 778
Db 1561 LLDPLAMDNHYGTQIPKEEMKSOEKSPEKTAFFKKDITLISLNACESHNAIAINEGNKP 1620
OY 779 819
Db 1621 EIEVTWAKQERTERLCSONPVLKRRHOREITRTTQSQOEEDYDQTSIWEKKKKEDPDY 1680
OY 820 DEDENQSPRSFOKTRHFIAAVERLMDYGMSSPHVLRNRAQSGSVPOFKVYFQEFTE 879
Db 1681 DEDENQSPRSFOKTRHFIAAVERLMDYGMSSPHVLRNRAQSGSVPOFKVYFQEFTE 1740
OY 880 GSFTQPLXRGELNEHLGLGPYIRAEVEDNIMVTFRNOASRPYSFSSLISYEEDQROGA 939
Db 1741 GSFTQPLXRGELNEHLGLGPYIRAEVEDNIMVTFRNOASRPYSFSSLISYEEDQROGA 1800
OY 940 EPRKRFVKNETKTYFMKVQOHMAPTKDEPDCKANAYFSVDVLEKDVASGLGPLVCHT 999
Db 1801 EPRKRFVKNETKTYFMKVQOHMAPTKDEPDCKANAYFSVDVLEKDVASGLGPLVCHT 1860
OY 1000 NTLNPAHGRQYTVQEFALFTIFDETksMYFTENNERNCRAPCNIOMDPTEKERYRHA 1059
Db 1861 NTLNPAHGRQYTVQEFALFTIFDETksMYFTENNERNCRAPCNIOMDPTEKERYRHA 1920
OY 1060 INGYIMDTLPGILWAQDORIRMYLISWGSNENIHSIHFSGHVTFYRKKEEKYKALYNLP 1119
Db 1921 INGYIMDTLPGILWAQDORIRMYLISWGSNENIHSIHFSGHVTFYRKKEEKYKALYNLP 1980
OY 1120 GVFEFVEMLPKAGIWRRECLIGEHLHAGMSTLPLVYNKOTPLGMASSGHIRDQRTAS 1179
Db 1981 GVFEFVEMLPKAGIWRRECLIGEHLHAGMSTLPLVYNKOTPLGMASSGHIRDQRTAS 2040
OY 1180 GOYGOMAPKLARLHYSGSINAMSTKEPFSWIKVDLAPMIIHGIKTQCARQKFSLSYISQ 1239
Db 2041 GOYGOMAPKLARLHYSGSINAMSTKEPFSWIKVDLAPMIIHGIKTQCARQKFSLSYISQ 2100
OY 1240 FIIMYSLDGKKWQTYRGNSTGTLMVFEQNVDSGIIKANIENPPIIARYIRLAPHTYSIRS 1299
Db 2101 FIIMYSLDGKKWQTYRGNSTGTLMVFEQNVDSGIIKANIENPPIIARYIRLAPHTYSIRS 2160
```



QY 1300 TLRMELMCCDLNCSGMPICMESKAISDAQITASSYFTNNFATWSPSKARLHLOGRSNMR 1359  
| | | | |  
Db 2161 TLRMELMCCDLNCSGMPICMESKAISDAQITASSYFTNNFATWSPSKARLHLOGRSNMR 2220  
| | | | |  
QY 1360 POWNPKMELQVDFQKTMKVTVGTQGVKSILTSMYKRELLSSSDGQHWTLFPONGKV 1419  
| | | | |  
Db 2221 POWNPKMELQVDFQKTMKVTVGTQGVKSILTSMYKRELLSSSDGQHWTLFPONGKV 2280  
| | | | |  
QY 1420 KVFQGNQDSFTFVNVNSLDPPLLTRYLRIHPQSMVHQIALRMEVLGCEADLY 1471  
| | | | |  
Db 2281 KVFQGNQDSFTFVNVNSLDPPLLTRYLRIHPQSMVHQIALRMEVLGCEADLY 2332  
| | | | |  
RESULT 15  
PCT-US93-03275-4.  
; Sequence 4, Application PC/TUS9303275  
; GENERAL INFORMATION:  
; APPLICANT: Lollar, John S.  
; APPLICANT: Runge, Marshall S.  
; TITLE OF INVENTION: Hybrid Human/Porcine Factor VIII  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Kilpatrick & Cody  
; STREET: 1100 Peachtree Street, Suite 2800  
; CITY: Atlanta  
; STATE: Georgia  
; COUNTRY: US  
; ZIP: 30309-4530  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US93/03275  
; FILING DATE: 19930407  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/864004  
; FILING DATE: 07-APR-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Padst, Patrea L.  
; REGISTRATION NUMBER: 31,284  
; REFERENCE/DOCKET NUMBER: EMO 106PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 404-815-6508  
; TELEFAX: 404-815-6555  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2332 amino acids  
; TYPE: AMINO ACID  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHEICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: N-terminal  
; ORIGINAL SOURCE:  
; ORGANISM: Homo sapien  
; TISSUE TYPE: Liver cDNA sequence  
; PCT-US93-03275-4  
Query Match 93.0%; Score 7315; DB 5; Length 2332;  
Best Local Similarity 62.3%; Pred. No. 0;  
Matches 1452; Conservative 0; Mismatches 0; Indels 880; Gaps 1;  
QY 20 ATRRYLGAVALSDMYMOSDLGELPYDARFPVRPKSPFNTSVVKKTLFVEFTVHLFN 79  
| | | | |  
Db 1 ATRRYLGAVALSDMYMOSDLGELPYDARFPVRPKSPFNTSVVKKTLFVEFTVHLFN 60  
| | | | |  
QY 80 IAKPPPMWGLGPTIOAEVYDTVVITLKNMASHPVSLHAGVSYWKASGAEYDDQTSO 139  
| | | | |

Db 61 IAKPPPMWGLGPTIOAEVYDTVVITLKNMASHPVSLHAGVSYWKASGAEYDDQTSO 120  
| | | | |  
QY 140 REKEDDKVFPGSGSHYVWQVLKENGPMASDPLCTYSLSHVLYKDLNSGLIGALLVCR 199  
| | | | |  
Db 121 REKEDDKVFPGSGSHYVWQVLKENGPMASDPLCTYSLSHVLYKDLNSGLIGALLVCR 180  
| | | | |  
QY 200 EGSIAKEKTQTLHKFTLLFAVDEGKSHSEFTKNSLMODROAASRAMPKHTVNGYNR 259  
| | | | |  
Db 181 EGSIAKEKTQTLHKFTLLFAVDEGKSHSEFTKNSLMODROAASRAMPKHTVNGYNR 240  
| | | | |  
QY 260 SLPLGLIGHCRKSYVMHVIQMGTPPEVHSIFLEGHFTLVNRHQASLESTPTFTLAOTL 319  
| | | | |  
Db 241 SLPLGLIGHCRKSYVMHVIQMGTPPEVHSIFLEGHFTLVNRHQASLESTPTFTLAOTL 300  
| | | | |  
QY 320 MDLGFLEFCHISSHQHDGMEAYVYVDSCEPEPQLRMKNNEAEEDYDDLTDSEMDVVR 379  
| | | | |  
Db 301 MDLGFLEFCHISSHQHDGMEAYVYVDSCEPEPQLRMKNNEAEEDYDDLTDSEMDVVR 360  
| | | | |  
QY 380 DDNSPSTQIRSVAKKHPKTVWHYIAEEDMDYAPLVLPADDRSYKSYLLNNGPORIG 439  
| | | | |  
Db 361 DDNSPSTQIRSVAKKHPKTVWHYIAEEDMDYAPLVLPADDRSYKSYLLNNGPORIG 420  
| | | | |  
QY 440 RYKRVKRFMAVYDEFTKREAIQHSGLIGPLLYGEGDFTLLIFKNASRPYNTYPRGI 499  
| | | | |  
Db 421 RYKRVKRFMAVYDEFTKREAIQHSGLIGPLLYGEGDFTLLIFKNASRPYNTYPRGI 480  
| | | | |  
QY 500 TVRPLYSRRLPKGVKHLKDFPLPGELFKYKWTVEGDPKSDPRCLTRYSSFVME 559  
| | | | |  
Db 481 TVRPLYSRRLPKGVKHLKDFPLPGELFKYKWTVEGDPKSDPRCLTRYSSFVME 540  
| | | | |  
QY 560 RDIASGLIPLLICYKESVDQGNQIMSDKNVILFSVDENRSWYLTENIORPLPNAG 619  
| | | | |  
Db 541 RDIASGLIPLLICYKESVDQGNQIMSDKNVILFSVDENRSWYLTENIORPLPNAG 600  
| | | | |  
QY 620 VOLEDEPQASNMHSINGVDFDSIQLSVCLHEVYWTLSGAOTDFLSYFSGSYTRKH 679  
| | | | |  
Db 601 VOLEDEPQASNMHSINGVDFDSIQLSVCLHEVYWTLSGAOTDFLSYFSGSYTRKH 660  
| | | | |  
QY 680 KMYEDTLTLFPFSGEIVSMENPGLWILGCHNSDFNRGWTALLKYSSCDKNTGDYE 739  
| | | | |  
Db 661 KMYEDTLTLFPFSGEIVSMENPGLWILGCHNSDFNRGWTALLKYSSCDKNTGDYE 720  
| | | | |  
QY 740 DSYEDISATLKSNNAIIEFRSFSQNSRHPSTROKOPNAT 778  
| | | | |  
Db 721 DSYEDISATLKSNNAIIEFRSFSQNSRHPSTROKOPNAT 780  
| | | | |  
QY 779 778  
| | | | |  
Db 781 KIONVSSDMLMLROSPTPHGLSLSDIOEAKYETFSDDPSGAIDSNNSISEMTHFRPQ 840  
| | | | |  
QY 779 778  
| | | | |  
Db 841 LHHSGDMVFTPESGIQLRLNEKLGTTAATLKKIDFKVSTSNLSTIPSDNLAAGTDN 900  
| | | | |  
QY 779 778  
| | | | |  
Db 901 TSLGPPSPVHYVDSQDITLFGKKSSPLTESGCLSLSEENNDKLESGIAMSQESSW 960  
| | | | |  
QY 779 778  
| | | | |  
Db 961 GKNVSTESGRLEFKGRAHPALLTKDNALFKVVISLTKTKTSNNSATNRKTHIDGSL 1020  
| | | | |  
QY 779 778  
| | | | |  
Db 1021 LIENSPSVQNLIESDTEKKVTPLIHDMLMDKNATLRLHMSNKTTSKKNMEMVQOK 1080  
| | | | |  
QY 779 778  
| | | | |  
Db 1081 KEGPIPPAQNDPMSEFKMLFLPESARWIOPTHGKNSLNSGQSPKOLVSLGPEKSYEG 1140  
| | | | |  
QY 779 778  
| | | | |  
Db 1141 QNFLESKRNVVVGKGEFTKDVGLKENVPPSSRNLFVTLNDLHNENNTIOEKKIOEETEK 1200  
| | | | |



QY 779 ----- 778  
Db 1201 KETLIQENVLPQIHVTYGTKNEMKNLFLSTRQNVESYEGAYAPVLODRSLNDSTNR 1260  
QY 779 ----- 778  
Db 1261 TKKHTAHSKKGEEENLEGLOQTQIVEKYACTTRISPTNSQONFVYORSKRALKQFRL 1320  
QY 779 ----- 778  
Db 1321 PLEETELEKRIIVDDTSTQWSKNMKHLTPSTLQIDYNEKEGAIQSPLSDCLTRSHSI 1380  
QY 779 ----- 778  
Db 1381 PQANRSPPIAKVSSPSIRPIYLRLVLFQDNSSHLPAASYRKKDSGVQESSHFLQAKK 1440  
QY 779 ----- 778  
Db 1441 NNLSLAILTEMTGDQREVSLGTSATNSVTYKKVENTVLPKPDLPKTSKVELLPKVHI 1500  
QY 779 ----- 778  
Db 1501 YQKDLFPETETNSGSPGHLDLVEGSLLOGTEGAIKWNEANRPKVPELRVATESAKTPSK 1560  
QY 779 ----- 778  
Db 1561 LLDPLAMDNHYGTQIPKREMKSQEKSPKTAFFKKDTILSLNACESNHAIAINEGONKP 1620  
QY 779 -----PVLRKHQREIRTRTLQSDQDEIDYDDTISVEMKKEDFDIY 819  
Db 1621 ELEVWAKQGRTERLCSQNPVLRKHQREIRTRTLQSDQDEIDYDDTISVEMKKEDFDIY 1680  
QY 820 DEDENOSRSPROKTRRHFIYAVERLMDYGGSSPHVLRNRAQSGSVQPKVYFQERTD 879  
Db 1681 DEDENOSRSPROKTRRHFIYAVERLMDYGGSSPHVLRNRAQSGSVQPKVYFQERTD 1740  
QY 880 GSFTQPLRGELNEHGLGYPYIRAEVDNIMVTFERNQASRPYSFSSLSIYEEDOROGA 939  
Db 1741 GSFTQPLRGELNEHGLGYPYIRAEVDNIMVTFERNQASRPYSFSSLSIYEEDOROGA 1800  
QY 940 EPRKNFVAPNETKTYFWKVQHMAPTKDEFDCKAMAYFSDVDLEKDVHSGLIGPLVCHT 999  
Db 1801 EPRKNFVAPNETKTYFWKVQHMAPTKDEFDCKAMAYFSDVDLEKDVHSGLIGPLVCHT 1860  
QY 1000 NTLNPAHGRQVTOEFALFTIPEDETSKNVTFENNERCARPCNIQMEDPTEKENYRFHA 1059  
Db 1861 NTLNPAHGRQVTOEFALFTIPEDETSKNVTFENNERCARPCNIQMEDPTEKENYRFHA 1920  
QY 1060 INGYIMDTLPGLVMAODQIRIMYLYLSMGSNENIHSIHSGHVFYVRKKEEYKMALYNLYP 1119  
Db 1921 INGYIMDTLPGLVMAODQIRIMYLYLSMGSNENIHSIHSGHVFYVRKKEEYKMALYNLYP 1980  
QY 1120 GVFEVTEMLPSKAGIWRVECLIGELHLAGMSTLFLVYSNKCQTPLGMAASGHIRDFOITAS 1179  
Db 1981 GVFEVTEMLPSKAGIWRVECLIGELHLAGMSTLFLVYSNKCQTPLGMAASGHIRDFOITAS 2040  
QY 1180 GQYGMAMPKRLARLHYSGSINAMSTKEPESWIKVDLAPMIITHGIKTOGAROKFESSLYISQ 1239  
Db 2041 GQYGMAMPKRLARLHYSGSINAMSTKEPESWIKVDLAPMIITHGIKTOGAROKFESSLYISQ 2100  
QY 1240 FIIMYSLDGKKMOTYRGNSTGTLMVFFGNVDSGIIKHNIENPILARIYRLHPTHYSIRS 1299  
Db 2101 FIIMYSLDGKKMOTYRGNSTGTLMVFFGNVDSGIIKHNIENPILARIYRLHPTHYSIRS 2160  
QY 1300 TLRRELKMGCDINSCSMPLGMSKASISDAQITASSYFTNMEATWSPSKARLHLQGRSNAMR 1359  
Db 2161 TLRRELKMGCDINSCSMPLGMSKASISDAQITASSYFTNMEATWSPSKARLHLQGRSNAMR 2220  
QY 1360 POVNNPKEMLOVDFOKTMKVGTTOGVKSLTSMYVKEFLISSQDGHQWTLFFONGKV 1419  
Db 2221 POVNNPKEMLOVDFOKTMKVGTTOGVKSLTSMYVKEFLISSQDGHQWTLFFONGKV 2280

QY 1420 KVFOGNDSEFTPVVNSLDPPLTRYLRIHPQSWVHQAIALRMEVLCGEAODLY 1471  
Db 2281 KVFOGNDSEFTPVVNSLDPPLTRYLRIHPQSWVHQAIALRMEVLCGEAODLY 2332

Search completed: November 17, 2001, 13:24:06  
Job time: 115 sec



**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 18, 2001, 04:58:30 : Search time 389.82 Seconds  
(without alignments)  
2768.374 Million cell updates/sec

Title: US-09-689-430-1\_COPY\_150\_4914

Perfect score: 4765

Sequence: 1 cttcttcaagtaacagta.....ttgggtcgttgcgcgcgc 4765

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued\_Patents\_NA:\*

- 1: /cgn2\_6/prodata/1/ina/5A\_COMB.seq:\*
- 2: /cgn2\_6/prodata/1/ina/5B\_COMB.seq:\*
- 3: /cgn2\_6/prodata/1/ina/6A\_COMB.seq:\*
- 4: /cgn2\_6/prodata/1/ina/6B\_COMB.seq:\*
- 5: /cgn2\_6/prodata/1/ina/PCRD\_COMB.seq:\*
- 6: /cgn2\_6/prodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4413	92.6	4629	2	US-08-484-891-7
2	4411.6	92.6	9354	1	US-08-683-8398-2
3	4357.6	91.5	4670	3	US-08-717-294-41
4	4330.8	90.9	4999	4	US-09-470-618-14
5	4330.8	90.9	4999	4	US-09-364-862-14
6	4259.4	89.4	11933	4	US-09-470-618-13
7	4259.4	89.4	11933	4	US-09-364-862-13
8	3748	78.7	5035	2	US-08-882-083-1
9	3748	78.7	5035	2	US-08-558-107-1
10	3748	78.7	5035	3	US-09-243-539-1
11	3189.8	66.9	4334	2	US-08-670-707A-38
12	3189.8	66.9	4334	4	US-09-037-601-38
13	2605.4	54.7	4451	3	US-08-717-294-42
14	2336.4	49.0	9009	1	US-07-864-004B-3
15	2336.4	49.0	9009	1	US-08-251-937A-3
16	2336.4	49.0	9009	1	US-08-212-133A-1
17	2336.4	49.0	9009	1	US-08-474-503-1
18	2336.4	49.0	9009	2	US-08-670-707A-1
19	2336.4	49.0	9009	4	US-09-037-601-1
20	2336.4	49.0	9009	5	PCT-US93-03275-3
21	2336.4	49.0	9009	5	PCT-US94-13200-1
22	2335.4	49.0	7056	1	US-08-121-202-1
23	2334.8	49.0	8241	6	5171844-1
24	2333.2	49.0	8967	1	US-08-366-851A-1
25	2275.2	47.7	6999	1	US-08-276-594A-1
26	1848	38.8	7032	4	US-09-324-867-1
27	1802.8	37.8	7493	1	US-08-212-133A-7

28	1802.8	37.8	7493	1	US-08-474-503-5	Sequence 5, Appl1
29	1802.8	37.8	7493	2	US-08-670-707A-5	Sequence 5, Appl1
30	1802.8	37.8	7493	4	US-09-037-601-5	Sequence 5, Appl1
31	1802.8	37.8	7493	5	PCT-US94-13200-5	Sequence 5, Appl1
32	1751.2	36.8	6402	2	US-08-670-707A-36	Sequence 36, Appl1
33	1751.2	36.8	6402	4	US-09-037-601-36	Sequence 36, Appl1
34	1240.4	26.0	1623	1	US-08-121-202-3	Sequence 3, Appl1
35	893.4	18.7	1130	1	US-07-864-004B-1	Sequence 1, Appl1
36	893.4	18.7	1130	1	US-08-251-937A-1	Sequence 1, Appl1
37	893.4	18.7	1130	1	US-08-212-133A-5	Sequence 3, Appl1
38	893.4	18.7	1130	2	US-08-474-503-3	Sequence 3, Appl1
39	893.4	18.7	1130	2	US-08-670-707A-3	Sequence 3, Appl1
40	893.4	18.7	1130	4	US-09-037-601-3	Sequence 3, Appl1
41	893.4	18.7	1130	5	PCT-US93-03275-1	Sequence 1, Appl1
42	893.4	18.7	1130	5	PCT-US94-13200-3	Sequence 1, Appl1
43	301.8	6.3	6909	2	US-08-804-196-1	Sequence 1, Appl1
44	301.8	6.3	6909	2	US-08-658-340-1	Sequence 1, Appl1
45	301.8	6.3	6909	3	US-08-746-111-26	Sequence 26, Appl1

## ALIGNMENTS

RESULT 1  
US-08-484-891-7  
Sequence 7, Application US/08484891  
Patent No. 5935935  
GENERAL INFORMATION:  
APPLICANT: Connolly, Sheila  
APPLICANT: Kaleko, Michael  
APPLICANT: Smith, Theodore  
TITLE OF INVENTION: Adenoviral Vectors for  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cecchi, Stewart & Olstein  
STREET: 6 Becker Farm Road  
CITY: Roseland  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07068  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch diskette  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WordPerfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/484,891  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/218,335  
FILING DATE: 25-MAR-1994  
APPLICATION NUMBER: 08/074,920  
FILING DATE: 10-JUN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Olstein, Elliot M.  
REGISTRATION NUMBER: 24,025  
REFERENCE/DOCKET NUMBER: 271010-273  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-994-1700  
TELEFAX: 201-994-1744  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4629 bases  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA primer  
FEATURE:  
NAME/KEY: Factor VIII cDNA with  
NAME/KEY: B domain deleted



US-08-484-891-7

Query Match 92.6%; Score 4413; DB 2; Length 4629;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 4419; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 271 atgcaaatagagctctccacgtctcttctctgctcttctgagatctctgctttagtgc 330  
DB 1 ATGCAAAATAGAGCTCTCCACCTGGCTTCTGTCCTTTGGCGATTCGCTTTAGTGC 60  
QY 331 accgaagatactactccggtggtgcagtggaactgtcatatgagactatagtcaaatgtatctc 390  
DB 61 ACAAGAAAGATACCTACCTGGGTGACAGTGAAGTATGATGGATATATGCAAGTATGTC 120  
QY 391 ggttagctgcctgtggaagcaagattccctctagagtgccaaatctcttccatcaac 450  
DB 121 GGTGAGCTGCTGTGGAGCGAAGATTCTCTCTAGAGTGCCAAATCTTTCCATTCAAC 180  
QY 451 acctcagtcgltgtaaaaaaagactctgttgttagaatcaaggttaccctttcaacatc 510  
DB 181 ACCTCACTCGTGTACAAAAGACTCTGTTGTAGAAATTCAGGTTACACCTTTTCAACATC 240  
QY 511 gctaaagccaagggccaacctgtgagtgtctgtctaggttccatccaaagctgaggttcat 570  
DB 241 GCTAAGCCAAAGGCCACCTGGATGGGTCTAGGTCTTACCATCCAGGCTGAGGTTTAT 300  
QY 571 gatcaagtggtcatctacacttaagaacaatgtctccatccctgtcagcttcatgtctt 630  
DB 301 GATACACTGGTCATTACACTTAAGAACATGGCTTCCCTCTGTCACTCTTCTATCTCTGT 360  
QY 631 ggtgtatctacttgaaaagctctctgaggagctgaatatgatgatcagaaccagtcaaaag 690  
DB 361 GGGTATCTCTACTGAAAAGCTTCTGAGGAGCTGAAATGATGATCAGACCAAGTCAAAAG 420  
QY 691 gaagaaagaatgataagttctccctgtgtggaagccaataatgcttggcaggtcctg 750  
DB 421 GAGAAAGAAAGTGAATGAAGTCTCCCTGGTGAAGCCATGATATGTTGGAGGTCCTG 480  
QY 751 aaagagaatgttccaatggtcctctgcaaccagtgtgcttacctactcatcttctcat 810  
DB 481 AAGAGATGCTCCAAATGGCTTGACCCACTGTGGCTTACTACTCTATCTTCTCTAT 540  
QY 811 gtggaactgtglaaaagactgtgaattcaggcctcatctgagccctactagatgtagaaga 870  
DB 541 GTGGACCTGTGAAAAAGACTTGAATTCAGGCTCTCATTTGGAGCCCTACTCTATGTGAGAA 600  
QY 871 gggagctgtgcgaagaaagacaacagaccttgacaacttatactactcttggctga 930  
DB 601 GGGAGTGTGGCAAGAAAGAACACAGACTTGGCAACAATTTATATCTTTTGTCTGTA 660  
QY 931 ttgtatgaaggaanaagttggaactcagaanaaagaactccttgatgcaagatagggat 990  
DB 661 TTGTGATGAAGGAAAGTTGGCACTCAGAAACAAAGAACTCCTTGATGTCAGAGATGGGAT 720  
QY 991 gctgcatctgtctcggtcgtgacctaaatgtcacacagtcgaatggttataaacaagtcct 1050  
DB 721 GCTGCACTCTGCTGGGCTGGCTGCTAAATGCACACAGTCAATGGTTATGTAAACGCTCT 780  
QY 1051 ctgcaggtctgtatgtatgccaaggaatcagttctatgtgcatgtgattggaaatgggc 1110  
DB 781 CTGCGAGGTGTGATGTGATGCGACAGGAATCACTATTTGGCATGTGATGGAAATGGGC 840  
QY 1111 accaactctgaatgtcaactcaatatctcctgaaggtcacacatctctgtgaggaacat 1170  
DB 841 ACCACTCTGAAGTGAAGTCAATATTCCTCGAAGTGCACATATTTCTTGTAAGAGAACAT 900  
QY 1171 cgcgaagcgtctcttggaatctcgccaataaacttctccttactgctcaaacactcttgatg 1230  
DB 901 CGCGAGCGCTCTTGAAGAACTCGCCCAATAACTTCTTACTGTGCAAAACACTCTTGATG 960  
QY 1231 gacctgggaacagttctcactgcttctgcatatactcttcccaacaaatgagatggaa 1290  
DB 1231 gacctgggaacagttctcactgcttctgcatatactcttcccaacaaatgagatggaa 1290

DB 961 GACCTTGGACAGCTTTCTACTGTTTTTGTCAATCTCTCCACCAACATGATGCGATGAA 1020  
QY 1291 gctatgtcacaagtagacagctgttccagaaggaccaccaactacgaatgaaataatgaa 1350  
DB 1021 GCTTATGTGCAAAAGTATGACAGCTGTCCAGAGAACCCCACTGACATGAAAAATGAA 1080  
QY 1351 gaagcgaagacatgatgatgatcttactgactctgaaatgaaatgtgtcaggttgat 1410  
DB 1081 GAAGCGGAAGACTATGATGATGATCTTACTGATTCGTAATGATGATGTCAGGTTTGAT 1140  
QY 1411 gatgaacactctctctcttcttataccaatctgcctcagttgtccaagaagcctccaaact 1470  
DB 1141 GATGACAACTCTCTCTTATCCAAATTCGCTCAGTGGCCAAAGACATCTTAAACT 1200  
QY 1471 tgggtacatacatgtcgtcgtgaagaggaagcgtgggacgtatgctccctagtcctgc 1530  
DB 1201 TGGGTACATTTACATTTGCTGCTGAAGAGAGACCTGGGACTATGCTCCCTTATGCTCGCC 1260  
QY 1531 ccggaatgaagaagttataaaagttcaatattgaaacaatggtccctcagcggatgtgtag 1590  
DB 1261 CCCGATGACAGAAAGTTATAAAGTCAATATTGAAACAATGGCCCTCAGCGATGTGAGG 1320  
QY 1591 aagtaaaaaaagttccgattatgtgctacacagatgaatcctttaaagctgtgaaact 1650  
DB 1321 AAGTACAAAAAGTCCGATTTATGGCATACACATGATAAACCTTTAAGACTGTGAAGCT 1380  
QY 1651 attcagcatgatacgaagactcttggaacttacttatatgaggagttggaacacatg 1710  
DB 1381 ATTGACGATGATCAGAAATCTTGGAGCTTTACTTTATGGGGAAGTTGGAGACACACTG 1440  
QY 1711 ttgattatattaaagatcaagcaagcagacalataacatctacccctcaggaatcact 1770  
DB 1441 TTGATTAATTTAAGAAATCAAGCAAGCAGACATTAACATCTCCCTCAGCATCTCT 1500  
QY 1771 gatgtcgtctcttgtatctcaagagattaccaaaagtgtaaaacttgaagatctt 1830  
DB 1501 GATGTCGCTCTTGTATTAAGAGATTAACAAAGGTGAATAACATTTGAAGGATTTT 1560  
QY 1831 ccaattctgcgaagaaatactcaaatataaagagcagtgactgtagaagatgagcca 1890  
DB 1561 CCAATTTGCGCAGAGGAATTTCAAAATTAATGAAGACAGTGTGAAGATGGGCCA 1620  
QY 1891 actaaatcagatcctcgtgctcgtgacccgctatactctagttcgttataatgagaga 1950  
DB 1621 ACTAATTCAGATCTCGGTGCTGACCCGCTATTACTCTGATTTCTGTAATGAGAGAA 1680  
QY 1951 gatctagcttcaagacatctgtgcccctctcctaactctgctacaaaagaactctgtagatcaa 2010  
DB 1681 GATCTAGCTTCAAGACTCATTTGGCCCTCTCTCATCTGTGTCACAAAGATCTGTAGATCAA 1740  
QY 2011 agaggaaccagataatgttcagacaagaagatgtcatcctgttctctgtattgtatgag 2070  
DB 1741 AGAGGAAMCCAGATTAATGTGTCAGACAAGAGAAATGTCAATCGTTTCTGTATTTGATGAG 1800  
QY 2071 aacggaagctgtgtactcctcaagagaataatacaacgcttcttcccactccagctgaggtg 2130  
DB 1801 AACCGAAGCTGTACCTCACAGAAATATACAGCGCTTCTCCCAATCCAGCTGAGAGTG 1860  
QY 2131 cagcttgaagatctcaaggttccaaagcctccaaactaaatgaagacatcaatgtgctatgt 2190  
DB 1861 CAGCTTGAGGATCCAGATTCAGATTCAGCTCCAAACTCATATGCAACATCAATGAGCTATGTT 1920  
QY 2191 ttgtatagtctgacgtgtgtcagttgttgtcataggtgtgcatatgtatctcaagc 2250  
DB 1921 TTTGATATTTGAGTGTGCTGATTTGTTGTGATGAAGTGGCATATGCTGATTAAC 1980  
QY 2251 atgtgagcagacgtgacttcttctgtcttctctctctgtgatatatccttcaaacacaa 2310  
DB 1981 ATTTGAGACAGACTGACTTCTTCTGTCTTCTTCTGTGATATACCTCAAAACACAAA 2040  
QY 2311 atgtgtatagaaacacatccactatcccaatctcaagagaactgtcttcatgtg 2370  
DB 2041 ATGGTATATGAAGACACTCACCTTATTCCTCATTTCCAGAGAACTGCTTATGTCG 2100



2371 atggaaccagctatgatcttgggtgcccacactcagacttcggaacagagc 2430  
2101 ATGGAACACCAGGTCTATGATTTCTGGGTGCACACACTCAGACTTTCCGAAACAGAGC 2160  
2431 atgacgcctactgaaggttctcagttgtgcaagaacacactggtgta ttaacgaagac 2490  
2161 ATGACGCCCTTACTGAAAGTGTTCAGTTGTGACAGAAACACTGGTGTATTTACGAGGAC 2220  
2491 agttatgaagatattcagacatacttgcctgagtaaaaaaaatgcacattgaaccaagaagc 2550  
2221 AGTTATGAAATATTTTCAGATCTTGCTGAGTAAAAAACATGCGCATTTGACACGAAGAGC 2280  
2551 ttctccagaattccaagaacccctagacactaggcaaaagaatttaattgcacccccca 2610  
2281 TTCTCCAGAAATTCAGACACCCCTAGCAGCTAGGCAAAAGCAATTTATAGCCACCCACCA 2340  
2611 gctctgaacacgcatcaacagggaaataactcgtactactcttcagtcaga tcaagagaa 2670  
2341 GCTTGAAAGCCCATCAACGGAATTAACCTGTACTACTCTTCAGTCAAGTCAAGAGAA 2400  
2671 attgactatgaatatacatacatagttgaaaatgaagaagaatttgacattatgat 2730  
2401 ATTGACTATGATGATACCATATCAGTTGAAATGAAAGAGAAATTTGACATTTATGAT 2460  
2731 gaagatgaataatcagaagcccccagcacttcaaaagaacacagacataattatgct 2790  
2461 GAGGATGAATAATAGAGCCCCCGAGCTTTCANAAAGAAACAGACCTATTTATTTGCT 2520  
2791 gcaatgagaagcgtcttgggata ttaggaatgagtagtccccacatgcttcaagaacag 2850  
2521 GCATGGAGAGGCTGTGGGATTTATGGGATGATAGCTCCACACATGTTCAAGAAACAG 2580  
2851 gctcagaatgagcagttccctcagttcaagaagaatggttttccaagaattactgagtc 2910  
2581 GCTCAGAGTGGAGTGTCCCTCAGTTCAAGAAAGTTGTTTCCAGGAATTTACTGATGGC 2640  
2911 tcttctactcagccctataacgltgagaacttaataagaacttggagacccctggggcca 2970  
2641 TCCTTTACTAGCCCTTATATCCGTGGAACTTAATGAATGAAATTTGGAGCTCCGGGCCA 2700  
2971 tataaagaacagaagttagaataataatcatgtaacttctcagaatcagagctcgt 3030  
2701 TATATAGACAGACAGATGTAAGATTAATCATGTGTAACCTTCAGAAATCAGGCTCTCGT 2760  
3031 cccattctcttacttactagccttatttctatgagaagaatcagaagacagaagca 3090  
2761 CCTATTCCTTCTATCTACGCTTATTTCTTATGAGGAAAGATCAGAGGCAAGGACAGAA 2820  
3091 cctagaaaaaacttctcaagcctaatgaagaacaaaacttacttggaaagtgcacat 3150  
2821 CTTAGAAAAAACTTTGTCAAGCCCTAATGANAACNAAACTTACTTTTGGAAAGTGCACAT 2880  
3151 cataatgacccaataaagaatgagttgactgcaagcctgggacttaattctctgagtc 3210  
2881 CATATGGCACCCACTAAAGATGAGTTGACTGCAAGCCTGGGCTTATTTCTCGATGTT 2940  
3211 gaactggaagaagaatgtgcaactcagagcctgaattggagcccttctgtctgcacaactaac 3270  
2941 GACTGGAAAAAAGATGTGCACTCAGGCTGATGGAACCCCTTCGTGGTCTCCACACTAAC 3000  
3271 acactgaacctgtcatgaggagaagaatgacagtaagaagaatttggctgtttttacac 3330  
3001 ACACGTGAACCTCTCATGTGGAGACAAAGTACAGTACAGGAATTTGCTCTGTTTTTCACC 3060  
3331 atcttctgatagaacaaaactgtgacttcaactgaaaata tggaaagaacactcagagct 3390  
3061 ATCTTTGATGAGACAAAAGCTGTGTACTCACTGAAAAATATGAAAGAAACTCGAGGCT 3120  
3391 cccctgcaataatcagaatgagaatcccaactttaaagaagaattatcgcttcagatc 3450  
3121 CCCTGCATATTCAGATGAGATGCCACTTTTAAAGAGAAATATGCTTCATGCAATC 3180

3451 aatgctacataatgataacactacccctgcttaagtaatggctcagaatcaagaattcga 3510  
3181 AATGGCTACATAATAGATATACACTACCTGCTTATATGCTCAGATCAAGATTTGCA 3240  
3511 tggatctgctcagcatlbggcaagcaatgaaaacatccattctatctatcttaattcgaat 3570  
3241 TGGTATCTGCTCAGCATGCGCAGCAATGAAAACATCCATTTCTATTTCAATTCAGTGACAT 3300  
3571 gttctcagctgaagaaaaaagagagataaaatggcactgtacactctctatccaggt 3630  
3301 GTGTCACTGTACAAAAAAGAGAGATATAAATATGACACTGATCAATCTCTATCCAGGT 3360  
3631 gttttgagaacagtggaagaattacacatccaaactgaaatttggcggtggaaatgctct 3690  
3361 GTTTTGAAGACAGGAAATGTTACATCCAAAGCTGGAATTTGGCGGTGGAAATGCTT 3420  
3691 atggcagacatcacatgctctggaatgagacacttcttctgtatcaagcaatgaatgt 3750  
3421 ATGGCAGCATCTACATGCTGGGATGAGCACACTTTTCTGCTGATCAGCAATTAAGTGT 3480  
3751 cagaatccctgggaatgagcttctggaacatataagaatttcaagatcaagctcagga 3810  
3481 CAGACTCCCTGGGAATGGCTTCTGACACATTAAGAATTTTCAGATTACAGCTTCAGGA 3540  
3811 caatatgacagtgaggcccaagaactggcagagacttcatattccggtacatcaatgctc 3870  
3541 CAATATGACAGTGGGCCCAAAAGCTGGCAGACTTATTTTCGGATCATATAATGCC 3600  
3871 tggagaccaagaagcccttcttggatcaaggttgagctgttctggaaccaatgatat 3930  
3601 TGGAGCACCAAGGAGCCCTTTTCTTGGATCAAGTGTGATGTTGGACCAATGATTTAT 3660  
3931 caaggacatcaagaacccagaggttgcctgcagaagttctccagcctctacatctcagttc 3990  
3661 CAGGCAATCAAGACCCAGGAGTCCCGTCAAGAAAGTTCTCCAGCCCTCATCTCCAGTTT 3720  
3991 atcatcatgatagtctcttgaatggaagaagtggagaacttaccagaagaattccacagga 4050  
3721 ATCATCTGATATAGTCTTGTGATGGGAAGAAAGTGCAGACTTATCGAGAAATTCACAGGA 3780  
4051 accttaatgctcttcttggcaatlgagatcatctlggataaaacaaatattttaa 4110  
3781 ACCTTATGCTCTTCTTGGCAATGTGATTCATCTGGATTAACCAATATTTTAAAC 3840  
4111 cctccaattatgctcgaatacatccgttgcacccaactcatatagcatctgcagcact 4170  
3841 CCTCCAAATTTATGCTCATCATCCGTTTGACCCCAACTCATTTATACATTCGACGACT 3900  
4171 ctccgcatgagttgagtgagcctgtgatttaataatgagttgcagcatgcatctggagaatgag 4230  
3901 CTTGCGATGAGTGTGATGGCTGTGATTTAAATAGTTTGACGATGCCATTTGGGAATGGAG 3960  
4231 agtaagaacatatacagaatgacagatctgcttcaacttcaacttcaacttcaacttcaact 4290  
3961 AGTAAGACATATACAGATGACAGATATGCTTACCTTACCTTACCAATATGTTGGC 4020  
4291 acctgtgtccttccaagaactcgaactcaccctcaagaaggagagtaatgcttggagact 4350  
4021 ACCTGTGCTCTCTTCAAAAGCTGCACTTCACTCCAAAGGAGAGATATGCTTGAAGACT 4080  
4351 caggtgataatccaaaagtgctgcaagtggaactccagaagaacatgaaagtcaca 4410  
4081 CAGTGAATATCCAAAAGAGTGGCTGCAAGTGCACCTCCAGAAAGCAATGAAAGTCCACA 4140  
4411 ggaatcaactcagaaggataaaatctctgtcttaccagcaatgataatgaaaggttctc 4470  
4141 GAGTAACTACTCAGGAGGATAAATCTCTCTTACAGCATGATGGAAGAGACTTCTC 4200  
4471 atctcagcagatcaagaatgagcatcagtgagctctcttcttcaagaatggaagtaag 4500  
4201 ATCTCCAGCAGTCAAGATGGCATCACTGAGCTCTCTTTTTCAGAAATGCAAGGTAAG 4260  
4531 gttttcagggaataacagactctcttcaacactgtgttgaactctctagcccaacgctta 4590



|||||  
Db 4261 GTTTTTCAGGAAATCAAGACTCTTTCACACCTGTGTGAACCTCTAGACCCACCGTTA 4320  
QY 4591 ctgactcgtcactctcgatctcaacccccagagttggtgacacagatgaccttgagatg 4650  
|||||  
Db 4321 CTGACTCGCTACTCTTGATTCACCCCAAGAGTTGGGTGCACCAATTTGCCCTGAGATG 4380  
QY 4651 gaagttctggtgctgagagcacaagagactctactgactcgagcgatgc 4699  
|||||  
Db 4381 GAGGTCTGGGCTGAGGACAGAGAGCTCTACTGAGGGTGGCCACTGC 4429

RESULT 2  
US-08-683-839B-2  
; Sequence 2, Application us/08683839B  
; Patent No. 5744326  
; GENERAL INFORMATION:  
; APPLICANT: ILL. Charles . R. et al.  
; TITLE OF INVENTION: Use of Viral Cis-Acting Post-Transcriptional  
; TITLE OF INVENTION: Regulatory Sequences To Increase Expression of  
; TITLE OF INVENTION: Intronsless Genes Containing Near-Consensus Splice Sites  
; NUMBER OF SEQUENCES: 18  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHYE & COCKFIELD  
; STREET: 60 State Street, suite 510  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109-1875  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/683,839B  
; FILING DATE: 11-MARCH-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Remillard, Jane E.  
; REGISTRATION NUMBER: 38,872  
; REFERENCE/DOCKET NUMBER: TTI-138  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617)227-7400  
; TELEFAX: (617)227-5941  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9354 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: CDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 2965..7378  
; US-08-683-839B-2

Query Match 92.6%; Score 4411.6; DB 1; Length 9354;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 4420; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 266 ccacatgcaaatagagctcaccgtctctctgctgaccttgcgattcgcttca 325  
|||||  
Db 2960 CCACCATGGTTTATGACTCTCCACCTGCTCTTCTGTGCTTTTGGATTCTGCTTTA 3019  
QY 326 gtgcacacagaagatactactggtgcaagtgcagctgcaatgagacataatgcaaatg 385  
|||||  
Db 3020 GTGCCACCAAGATCTACTGAGGTGAGTGCAGATGCACTATGCAAAAGTG 3079

QY 386 atctcgtgagctgctctgtagcgaagaattctcctctatagatgccaataatcttccat 445  
|||||  
Db 3080 ATCTCGTAGAGCTGCTGCTGAGCGCAAGATTTCTCTAGAGTGCAAAATCTTTTCCAT 3139  
QY 446 tcaaacactcagtcgtgtacaaaaaagactcgttltgtgaattcaacggttacctttca 505  
|||||  
Db 3140 TCACACCTCACTGCTGTACAAAAGACTCTGTTGTGTGAATTCACGGTTTCACTTTTCA 3199  
QY 506 acatcgtcaagccaagcgccacctgtagtctgtcgtatgctccaccaccacagctag 565  
Db 3200 ACATGCTAAGCCAGGCGCACCCCTTGATGGGTCTGTAGTCTTACCAATCCAGGCTGAGG 3259  
QY 566 ttatgatacagtggtgcatlaacacttaagaacatggtcttccactcgtgctatcgt 625  
Db 3260 TTTATGATACAGTGTGCTTACACTTAAGAACATGGCTTCCATCTGTGCTTTCATG 3319  
QY 626 ctggtggtgtatccactcgtggaagcttctggtggaagctggaatagatgatacgaagc 685  
Db 3320 CTGTTGGTGTATCTTACGTGAAAGCTTCTGAGGGAGCTGAATATGATGATCAGACAGTC 3379  
QY 686 aaagagagaagaagatataaagcttcctcctggtggaagcatatgctgagcag 745  
Db 3380 AAAGGAGAAAGAAATATATAAGTCTTCCCTGGTGGAGGCAATACATATGCTGCGCAGG 3439  
QY 746 tctgaaagagaatggtcacaatggtcctctgaaccactggtcctactactatctt 805  
Db 3440 TCCTGAAGAGATGTGCTCAATGGCTGTGACCCACTGTGCTTACCTACATATCTTT 3499  
QY 806 ctcatgtgacctggttaaaagacttgaattcaagcctcatcttgagccctactatgta 865  
Db 3500 CTCATGTGACCTGTGTTAAACATTTGAATTCAGGCTTCATTGGAGCCCTACTATATGTA 3559  
QY 866 gagaagggagctcgtgccaaggaagaagacagactctgccaataattactactttg 925  
Db 3560 GAGAAAGGAGCTGTGCGCCACAGGAAAAAGACACAGACTTGTCACAAATTTATCTACTTTTG 3619  
QY 926 ctgtattgtatgaagggaaaagtgtggcactcagaacaaagaactcctgtatgagata 985  
Db 3620 CTGTATTTGATGAAGGGAAGAGTTGGCACTCAGAAACAAAGAACTCTGTATGCAAGATA 3679  
QY 986 gggatgctgcatctgctcgtggtcgtcctaaatgacacagtaatggttatgtaaca 1045  
Db 3660 GGGATGCTGCATCTGCTGGGCTGGCCCTTAAAGACACAGTGAAGTTATGTTAACA 3739  
QY 1046 ggtctcgtccaggtctgattgtagtgcacagaagaatcaagctatcttggaatgtaaa 1105  
Db 3740 GGTCTCTGCAAGGTGTGATTGGATGCCACAGAAATCATGTTATGGCATGTGATTGAA 3799  
QY 1106 tgggacacactcctgaaagtgcaactcaatatctcctgaaaggtcacacattcctgtgaga 1165  
Db 3800 TGGGCACCACTCTGAAAGTGCACATCAATATTCCTGAAAGTCAACATTTCTGTGAGGA 3859  
QY 1166 accatcgccagggcgttcccttggaatatctgcacaataacttctactcgtccaacactc 1225  
Db 3860 ACCATCGCCAGGCGCTCTGGAATCTCGCCCAATATCTTCTTACGTGCAACACTCT 3919  
QY 1226 tgaatgaccttggagacgtttctactggttltgtcatctctcctccacaacatgagta 1285  
Db 3920 TGATGAGCTTGGACAGTCTTACTGTTTGTCTATCTCTTCCACCAACATGATGGA 3979  
QY 1286 tggaaagctatgtcaaaatagacagctgtccagagaagaccccaactagaatgaaaaata 1345  
Db 3980 TGGAAAGCTTATGTCAAAAGTAGACAGCTGTCCAGAGGAACCCCACTACGATGAAATA 4039  
QY 1346 atgaaagagcggaagactatgtagatgacttactgattctgaaatgtagtgtgcaagt 1405  
Db 4040 ATGAAGAGCGGAAGACTATGATGATGATCTTACTGATTCGAAATGAGATGTGTCAGGT 4099  
QY 1406 ttgatgtgacaactcctcctcttcaacaaattgctcgaagttgccaagaagacatccia 1465  
Db 4100 TTGATGATGACAACTCTCTCTTTTATCCAAATTCCTCAAGTTGCCCAAGACATCTTA 4159  
QY 1466 aaacttgggtacattacatltgctgtgaagagagagacttggagatagtcctcctagtc 1525



Db 4160 AACCTGGGTACATTACTCTGCTGAGAGAGAGACTGGACTATGCTCCCTAGTCC 4219  
QY 1526 tcgcccccgatgacagaagtataaagtcataattgacaatrgccctcagcgatgtg 1585  
Db 4220 TCGCCCCGAGAGAGAAGATTATAAAGTCAATATTTCACAAATAGGCCCTCAGGGGATGG 4279  
QY 1586 gtagaagttacaaaaaagctccgatttactgatacacagatgaacctttaaagctgtg 1645  
Db 4280 GTAGGAAGTACAAAAAGTCCGATTATTGCAATCACAGATGAAGACTTTTAGACTGCTG 4339  
QY 1646 aagctatcagcagtaacacagaaactctgagaccccttacttacttgagggaagcttgagaca 1705  
Db 4340 AAGCATATTCACATGAAATCAGGAATCTTGGACCTTACTTATGGGGAAGTGGAGACA 4399  
QY 1706 cactgttgatataatttaagaatacaagcagacagacataaactcaccctcagga 1785  
Db 4400 CACTGTGATTATATATTAAAGATCAAGCAACAGACCATTAACATCACTACCGGAA 4459  
QY 1766 tcactgagtcgcgccttctgtatctcaaggagattaccaaaagtgtaaaacattgaag 1825  
Db 4460 TCACGTGATGTCCTGCTCTTGTGATTCAAAGAGATTAACAAAAGTGTTAAACATTGGAAG 4519  
QY 1826 atttccaattctgcagaggaataatccaataataatgagacagtactgtagaagatg 1885  
Db 4520 ATTTTCAATTCTGCCAGAGAAATATTCAAAATATATGACAGTGCCTGACAGATG 4579  
QY 1886 ggccaactaaactcagactcctcgtgtgcctgacccgtactactagttcgtaatatgg 1945  
Db 4580 GGCCAACTAAATCAGATCCCTCGGCTGACCCGCTATTACTAGTTCCGTAAATAGG 4639  
QY 1946 aagaagcttagcttcaaggactcattggccctccctccatcgtgttaaagaactctag 2005  
Db 4640 AGAAGATCTAGCTTCAGGACTCATGTGACCTCTCTCATCTGCTACAAAGAACTGTAG 4699  
QY 2006 atcaaaaggagaacagataatgcaagaacagagaaatgcatcctgttctctgatttg 2065  
Db 4700 ATCAAGAGGAAACACAGATATATGTCAGACAAAGAGAAATGTCATCTGTATTTTG 4759  
QY 2066 atggaacccgaagctgtgtacactcagagaaatatacaagcttctcccaatccagctg 2125  
Db 4760 ATGGAACCCGAAAGCTGTACTCTCACAGAGAAATATCAACGTTTCTCTCCCAATCAGATG 4819  
QY 2126 gaggcagcttggaagatcccaagatcccaagctccaacatcagcagacataatgct 2185  
Db 4820 GAGTGCAGCTTGAGGATCCAGAGTTCCAAAGCTCCAAATATATACAGCAATCATGCT 4879  
QY 2186 atgttttgaatgttgcagttgtcagttgttgcataagagtggaactactgtacatc 2245  
Db 4880 ATGTTTGAATGTTTGCAGTGTGCAGTTGTTTGCATGAGGTGSCATACTGTAATTC 4939  
QY 2246 taagcattgagacagacagactccttctctgtctctctctctctcgatataactcacaac 2305  
Db 4940 TTAACATTGGAGCAACAGATGACTTCTTCTGTTCTTCTCTGATATACCTTTCAAC 4999  
QY 2306 acaaatggtctatagaagacacactcaccctcattccatctcaggggaacatcttcca 2365  
Db 5000 ACAAAATGGCTTATGAAGACACACTCACTCACTATTCCTCAATTCAGAGAAATCTCTTCA 5059  
QY 2366 tctcagatgaaaaaacccagatctatgatactctgggtgtgccaaactcagacttctggaaca 2425  
Db 5060 TGTGATGAGAAACCCAGGTCTATGATTTCTGGGGTGCACAACTCAGACTTTGGAAACA 5119  
QY 2426 gaggcagtcagccgcttactcgaagttctctagttgtgacaagaacaactggtgataatcag 2485  
Db 5120 GAGCATGACCGCCTTACTGTAAGGTTCTAGTTTGTGACAAACACATGTCATATTATACG 5179  
QY 2486 aggcagcttatgaagacatttcagcatactctgtagtaaaacaatgccaatgaaacaa 2545  
Db 5180 AGGACAGTTATGAAGATATTTACACATATCTTGCTGAGTAAAAACAATGCCATTGAACCA 5239  
QY 2546 gaagcttctccagaatccaagacacctagcacttagcaaaagcaatttaatgccacc 2605

Db 5240 GAAGCTTCTCCAGAAATTCAGAGACACCCCTAGCACTAGGCAAAAGCAATTTAATGCCACC 5299  
QY 2606 caccagctcttgaaaacgcacatcaacggaataaactcgttactactcttcaatcgaatcaag 2665  
Db 5300 CACGAGTCTTGAAGCGCATACCGGAAATACCTCGTACTACCTCTCAGTCAATCATCAG 5359  
QY 2666 aggaattgactatgatatccatcagatgtgaatgaagaagaagatttgcattt 2725  
Db 5360 AGGAATTTGACTATGATGATATCCATATCAGTTGAAATGAGAAAGAGATTTTGCATTT 5419  
QY 2726 atgataagatgaaatcgaagcccccgcagcttctcaaaagaagaacacagacataatla 2785  
Db 5420 ATGATGAGAGATGAATAATCAGAGCCCCGACCTTTCAAAAAGAAACAGACATATTTTA 5479  
QY 2786 ttgtctcagtgagaaagctctcgggattatgagatgagtagtctcccaatgttttaagaa 2845  
Db 5480 TTTGCTGAGTGAAGAGGCTCGGATTTATGGAGTAGTACCTCCCACTATCTTGAAGAA 5539  
QY 2846 acaaggctcagagtgagcagtgctccctcagttcagtaagaagttgtttccaggaattactg 2905  
Db 5540 ACAGGGCTCAGAGTGCAGTCCCTCAGTTCAAGAAAGTTGTTTCCAGGAATTTACTG 5599  
QY 2906 atggctccttactcagcccttataccgttgagaaactaaatgaaacatttggactcctg 2965  
Db 5600 ATGGCTCCTTACTCAGCCCTTATACCGTGAAGAACTTAATGAAACATTTGGAGACTCTGG 5659  
QY 2966 ggccatatataagacgaaagttagaataatcaatgtagtacttccaagaataagcct 3025  
Db 5660 GGCCATATATTAAGACGAAGAACTTAAGATATATCATGTGTAATCTTGAATCAGGCT 5719  
QY 3026 ctgctccctatctcttctcattcagcttacttcttgaagaagaatcagaggcagaag 3085  
Db 5720 CTGCTCCCTATTCCTTCTATTTCTAGCTTATTTCTTATGAGGAAGATCAGAGCAAGAG 5779  
QY 3086 cagaacccagaaaaaacttctgcaagccataatgaacccaaacttacttcttgaagatg 3145  
Db 5780 CAGAACCTGAGAAAAAATCTTGTCAAGCCTAATGAACCAAACTTCTTGGAAAGTGC 5839  
QY 3146 aacatcataatggcaacccataaagatgatttgcactgcaaaagccttggcttattctctg 3205  
Db 5840 AACATCATATGCGAACCCACATTAAGATGAGTTGACTGCAAAAGCCCTGGCTTATTTCTGTG 5899  
QY 3206 atgttgaacctgaaagaatgtgcactcagccggaattggaccccttctgtctgcaca 3265  
Db 5900 ATGTTGACCTGGAAGAAAGATGTGCATCAGGCTGATTTGACCCCTTGTGCTGCCACA 5959  
QY 3266 ctaacacacttgaacccctgtctcaatggagacaagtgcagatcaagaatttgcctgtttc 3325  
Db 5960 CTAAACACACTGAAACCTGCTCATGTGAGACAAAGTACAGTACAGGAATTTGCTGTTTT 6019  
QY 3326 tcaccatctttagatgagacccaaaagctgtgtacttccactcgaaatatgaaagaactgca 3385  
Db 6020 TCACCATCTTTGATGAGACCAAAAGCTGTACTTCACTGAAAAATATGGAAGAAACCTGCA 6079  
QY 3386 gggtccctgtcaatcagatggaatgccactttaaagaagaattatgccttccatg 3445  
Db 6080 GGGCTCCCTGCAATATTCAGATGGAAGATGCCACTTTTAAAGAAATATATGCTTCCATG 6139  
QY 3446 caatcaatggtctacataatgatacactcctgtgcttagtaatggtctcagaatcaagaaga 3505  
Db 6140 CAATCAATGCGCTACATATATGATGATACACTGCTGCTTATGTAATGGCTCAGGATCAAGGA 6199  
QY 3506 ttcgatgtagtctgtcgaacatggcagcaatgaaatcaatccattctatattcgaatg 3565  
Db 6200 TTTGATGGTATCTCTCAGCTGAGGCAAGCAATGGAACATCAATTCATTTCAATTTCACTG 6239  
QY 3566 gacatgtagtcaactgtaaaaaaagagagtaaaaaatgcaactgtacatctcctc 3625  
Db 6260 GACATGTGTTACCTGTAAGAAAAAGAGAGTAAATAATGCACTGTACAAATCTCTATC 6319  
QY 3626 caagtgattttgagacagtggaatgttaccatccaaagcttgaatttgcgggttggaat 3685  
Db 6320 CAGGTGTTTGAAGACAGTGAATGTTACCATCCAAAGCTGGAATTTGCGGGTGGAAAT 6379







|||||  
Db 444 AGGGAAGAAAGATGATGATAAAGTCTCCCTGGTGGAAAGCCATATATGCTGGCAGGTC 503  
Qy 748 ctgaagaagaatggtccaatggtccctctgaccacgtgcttactactatacttctc 807  
Db 504 CTGAAGAAGAAATGGTCCATATGGCCCTTGACCCACTGTCCTTACTCTCATATCTTTCT 563  
Qy 808 catgtggaccctggtlaaagaacttgatccagccctcatctggagccctctatgtaga 867  
Db 564 CATGTGGACCTGGTAAAGACTTGATTCAGGCCCTCATTTGGAGCCCTCATGTATGTGA 623  
Qy 868 gaaaggagctctggccaaggaaagacacagacctctgcaaaattatactacttltgct 927  
Db 624 GAAGGAGCTCTGGCCAAAGAAAGACACAGACCTTGACAAATTTATCTACTTTTCT 683  
Qy 928 gtatttgatgaaggaaagtctggaccctcagaacaagaactccctgaatgcagaatagg 987  
Db 664 GTATTGTGAAGGGAAGGTTGGCACTCAGAAACAAAGAACTCCTTGATGCAGATAGG 743  
Qy 988 gatctgcatctgctcgggcttggccttaaaatgcacacagtcacatggttatgaacagg 1047  
Db 744 GATGCTGATGTGCTCGGGCCTGGCCTTAATGCACACAGTCAATGGTTATGTAAACAGG 803  
Qy 1048 tctctgccaaggtctgatatggaatgcacaggaatcaatctctatggcatgtggaatg 1107  
Db 804 TCTCTGCCAGGCTGATGTGATGGCCACAGGAATCAGTCTATTGGCATGTATGGAAAG 863  
Qy 1108 ggcacacactcctgaagtgacatcaatctcctgaaggtcacacattctgtgaagaa 1167  
Db 864 GGCACCACTCTCTGAAGTGACCTCAATATCTCCAGAGTCCACACATTTCTTGAGGAAAC 923  
Qy 1168 catgcgacggcgctcttggaaatctcgccaataacttccctacttgctcaaaactctg 1227  
Db 924 CATGGCCAGGGCTCTTGGAATCTCCGCAATTACTTTCTTACTGCTCAACACTCTTG 983  
Qy 1228 atggaccttgagacagtttctactgttttgcatactctctccacacacatgtagatg 1287  
Db 984 ATGGACCTTGGACAGTTTCTACTGTTTGTGATATCTCTCCACCAACATGATGGCATG 1043  
Qy 1288 gaagacttatgcacaagtatgacagctgtccagaagaaacccaactcagaaatgaataat 1347  
Db 1044 GAACCTTATGTCAAAAGTAAAGACGTGTCCAGAGAACCCCAACTACGATGAAAAATAT 1103  
Qy 1348 gaagaagcggaagactatgacatcacttacttacttgaatgtagtgtgaagttt 1407  
Db 1104 GAAGAAGGGAAAGACTATGATGATGATCTTACTATTTCTGAATGTGATGTGAGGTTT 1163  
Qy 1408 gatgtgacaactcctccttcttcttataccaaatctgctcaagttgcagaagaacatctaa 1467  
Db 1164 GATGATGACACACTCTCTCTCTTATCCAAATTCGTCAGTTGCCAAGAAGCATCTTAA 1223  
Qy 1468 acttggtgacttatactgtcgtcgtgaagaaggaagctggtgactatgctccttactc 1527  
Db 1224 ACTTGGGTACATTTACATTTGCTGGAAGAGAGACTGGGACTATGTCTCTTGTGCTCTC 1283  
Qy 1528 gcccccgaatgacagaagttataaagaatcaatatttgaacaatggccctcagaagatggt 1587  
Db 1284 GCCCCGATGACAAAGTTATTAAGATCAATATTGGAACAATGGCCCTCACGGATTTGGT 1343  
Qy 1588 aggaagatcaaaaaagtcgaattatgtgcacacacagatgaacacttlaaagaccgtgaa 1647  
Db 1344 AGGAAGTAAACAAAAGTCCGATTTATGGCATACACAGATGAACCTTTAAGACTCGTGAA 1403  
Qy 1648 gctattcaagcattgaatcaagaaacttgggaacttacttctatggggaagtgtgaacaca 1707  
Db 1404 GCTATTGAGCTGATGATGAGATCTTGAGACTTTACTTTATGGGAATGGGAGACACA 1463  
Qy 1708 ctgttgattatatttaagaatcaagcaagcagacacataaacatctaccctcagaagatc 1767  
Db 1464 CTGTTGATTAATTAAAGATCAAGCAAGACCATTAATTAACCTTCACGGAAATC 1523  
Qy 1768 actgattcgcctctgttattcaagaagatctacaaagggtgaacaactltgaagat 1827  
|||||

Db 1524 ACTGATGTCCCTCTTCTTATTCAAGGAGATTACCAAAAAGGTAAACATTTGAAGAT 1583  
Qy 1828 ttccaattctgccaagggaatatltaaatataatgaaatgaaatgaaatgaaatgaaatg 1887  
Db 1584 TTTCCAAATTTGCCAGGAGGAATATTAATTAATTAATGAAGACGACTGTGAAGATGGG 1643  
Qy 1888 ccaactaaatcagaatccctgggtgcctgacccgacttacttacttcttcttgaatagag 1947  
Db 1644 CCAACTAATATAGATCCCTGGTGCCTGACCCGCTATTACTTACTTGTCTTAATATGGAG 1703  
Qy 1948 agagatctaagcttaaggactatctggccctctcctcatctgctacaaagaatctgtgat 2007  
Db 1704 AGAGATCTAGCTTCAGGACTCATTTGGCCCTCTCCATCTGCTCAAAAGAAATCTGTAGAT 1763  
Qy 2008 caaaggagaaacagataatgctcagaagaagaaatgcatcctglttctgtatttgat 2067  
Db 1764 CAAAGAGGAAACAGAAATATGTACAGAAAGGAATGCATCTGTTTGTATTTGAT 1823  
Qy 2068 gagaaccgaagctgttactcctacagagaatatataacgcttctcccaatccagctgga 2127  
Db 1824 GAGAACCGAAGCTGTGTACTTACAGAGATATACAGCTTTCTTCCCAATTCAGCTGGA 1883  
Qy 2128 gtgcagcttgaagaatccagagttccaaagctcacaacalcatgcaagaatcaatgctat 2187  
Db 1884 GTGCAGCTTGAGATGCCAGATTCCAAGCTCCAAACATCATGACAGATCAATAGGCTAT 1943  
Qy 2188 gtttttgaatgttgcagttgtcagttgttgcatagaagtgtgcatactgtlactta 2247  
Db 1944 GTTTTATATATTGTGCAAGTTGTCAAGTTGTTGATGAGGAGCATCTGTGATATTCTA 2003  
Qy 2248 agcatggagacagactgacttcttctgtcttcttcttcttcttgaatatacttcaaacac 2307  
Db 2004 AGCATTTGGACACAGACTGACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCAAC 2063  
Qy 2308 aaaaatgtctataagaacacactacactatcccatctcagaagaaactgttctcatg 2367  
Db 2064 AAATGTGCTATGAAGACACACTCACCTATTTCCATTTCCAGAGAAAGTCTTCATG 2123  
Qy 2368 tcatgtgaaaaacccaagtctctatgattctgtgggtgtgccaaactcagacttgcgaaca 2427  
Db 2124 TCGATGGAAGAAACCCAGGTCTATGATTTCTGGGGTGCCACAACTCAGACTTTCGGAAC 2183  
Qy 2428 ggcattgacccgcttacttgaagagtttctagttgtgaagaagaactgtgtatttactag 2487  
Db 2184 GGCATGACCCGCTTACTGAAAGTTTCTAGTTGTACAAAGAACCTGGTATTTATTCAG 2243  
Qy 2488 gacagttatgaagaatttcaagcaacttactgtctgaatgaagaacaaatgcatgtgaacca 2547  
Db 2244 GACAGTTATGAAGATATTTGACATTTGCTGTGATTAATAAACAATGGCATTTGAACCAAG 2303  
Qy 2548 agcttctcccaagaattcaagaacacctagaactagccaagaacaaatlaatgcccaca 2607  
Db 2304 AGCTTCTCCCAAGATTCAAGACACCTTAGCACTAGGCAAAAAGCAATTAATGCCCACCA 2363  
Qy 2608 ccagttctgaagaagccatcaacgggaataactgtactactcttcaagtaagatcaagag 2667  
Db 2364 CCACTCTTGAACCCCAATCAAGCGGAATATCTGTCTTCTTCAAGTCAAGTCAAGAG 2423  
Qy 2668 gaaattgactatgataacacatactcagttgaatgaagaagaagaatttgacattat 2727  
Db 2424 GAATTTGATATGATGATGATACATATCAGTTGAATGAAGAAGAAATTTGACATTTAT 2483  
Qy 2728 gatgaagatgaataatcagaagcccccgaagcttcaaaagaagaacagacatattatc 2787  
Db 2484 GATGAGATGAAATATGAGACCCCCGAGCTTCAAAAAGAAACACACACTATTTATTT 2543  
Qy 2788 gctgcattgagaagggcttggatattgagatagtagtccccaatggttctaaagaac 2847  
Db 2544 GCTGCATGAGAGGCTCTGGGATTTATGGATGATGACTCTCCACATGTTCTTAAGAAC 2603  
Qy 2848 agggctcagaagtgtccatgtccctcagttcagaagaagtgttctcagaatatacagat 2907  
Db 2604 AGGGCTCAGAGTGCGCACTGTCCTCACTTCAAGAAAGTTGTTTCCAGAAATTTACGAT 2663



2908 ggtcccttactcagcccttataccgttgagaactaaatgaacatttggaccttggg 2967  
2664 ggcctcttttactcagccctttatccgttgagacactaaatgaacatttggaccttggg 2723  
2968 ccatataaagaagaagtgtgaagataatcatctgttaacttcaagaatcaggctctc 3027  
2724 ccatatattatgaagcagaagattgaagatattcatagttacatttgcgaatcagccctct 2783  
3028 ggtcccttactcagcccttataccgttgagaactaaatgaacatttggaccttggg 3087  
2784 ggtcccttactcagcccttttactcagcccttttactcagcccttttactcagcccttt 2843  
3088 gaacctgaagaaacttctgaagccttaataaaccacaaacttacttcttgaagtgcaa 3147  
2844 gaacctgaagaaacttctgaagccttaataaaccacaaacttacttcttgaagtgcaa 2903  
3148 catcataatgagcaccacaaagatgtgactgtgacgtgcaagccttcttctctgac 3207  
2904 catcatattgagcaccacaaagatgtgactgtgacgtgcaagccttcttctctgac 2963  
3208 gtttacccttgagaagaaagtgtgactcagccttgatgtgaccccttctgtctgcaagc 3267  
2964 gtttacccttgagaagaaagtgtgactcagccttgatgtgaccccttctgtctgcaagc 3023  
3268 aacacactgaacccctgtctcaatgagagacaagtgtgacagagaaatttgcctgttctc 3327  
3024 aacacactgaacccctgtctcaatgagagacaagtgtgacagagaaatttgcctgttctc 3083  
3328 aacacttctgtatgagaccaaagctgtgactcactgtaaaataatgaaagaaactgtcag 3387  
3084 accattcttctgtatgagaccaaagctgtgactcactgtaaaataatgaaagaaactgtcag 3143  
3388 gttcccttgagaatatacagaatgagaatccacatttgaagaatatacagcttccatca 3447  
3144 gcttcccttgagaatatacagaatgagaatccacatttgaagaatatacagcttccatca 3203  
3448 atcaatgtctacataatgatacactacactgtgactgaatgagccagagatcaaaagatt 3507  
3204 atcaatgtctacataatgatacactacactgtgactgaatgagccagagatcaaaagatt 3263  
3508 cgaatgtatctgtcgaagacatgagcagaatgaataacatccatctcatctcaatgag 3567  
3264 cgaatgtatctgtcgaagacatgagcagaatgaataacatccatctcatctcaatgag 3323  
3568 catgtatctacatgatacagaagaaagagataaataatgagcctgtacaatctcatca 3627  
3324 catgtatctacatgatacagaagaaagagataaataatgagcctgtacaatctcatca 3383  
3628 ggttcttcttgagaagatgaaatgttacaatccaaagcttggaatttggcgggtggaatgc 3687  
3384 ggttcttcttgagaagatgaaatgttacaatccaaagcttggaatttggcgggtggaatgc 3443  
3688 ctatattgagacatctacatgctgagatgagacacatttctgtgtgtacagcaataag 3747  
3444 ctatattgagacatctacatgctgagatgagacacatttctgtgtgtgtacagcaataag 3503  
3748 tgtcagactcccttgagagatggtctctgacacatctagaatcttcagatcagacttca 3807  
3504 tgtcagactcccttgagagatggtctctgacacatctagaatcttcagatcagacttca 3563  
3808 ggaacataatgagacatgagcccaaaagcttgagcagacttcatctatctcagatcaat 3867  
3564 ggaacataatgagacatgagcccaaaagcttgagcagacttcatctatctcagatcaat 3623  
3868 ggtctgagacacaaagagcccttcttctgtatcaaggtgtgatctgtgtcagcaatgatt 3927  
3624 ggtctgagacacacaaagagcccttcttctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 3683  
3928 attcagagacacaaagacacaaaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 3987  
3684 attcagagacacacaaagacacaaaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 3743

3988 ttatcatcatgataagcttctgataaggagaagatgagcagacttaccaggaatctccat 4047  
3744 ttatcatcatgataagcttctgataaggagaagatgagcagacttaccaggaatctccat 3803  
4048 ggaacataatgagacatgagcccaaaagcttgagcagacttcatctatctcagatcaat 4107  
3804 ggaacataatgagacatgagcccaaaagcttgagcagacttcatctatctcagatcaat 3863  
4108 aacctccaatatgtctcgaatacaccgttgcacccaatctataagatctcagc 4167  
3864 aacctccaatatgtctcgaatacaccgttgcacccaatctataagatctcagc 3923  
4168 actcttcagatgagatgataaggcctgtgataatagatgtaagcagatgtaagatg 4227  
3924 actcttcagatgagatgataaggcctgtgataatagatgtaagcagatgtaagatg 3983  
4228 gagaataagcaatacagatgagcagatgtactgtcttcatcttaccatattt 4287  
3984 gagaataagcaatacagatgagcagatgtactgtcttcatcttaccatattt 4043  
4288 ggcacctgtctccttcaaaagctcagactcactcacaagagagatgtaagcctgaga 4347  
4044 ggcacctgtctccttcaaaagctcagactcactcacaagagagatgtaagcctgaga 4103  
4348 cctcagatgataatccaaagatgagcctgagatgagacttccagaagacatgaaatg 4407  
4104 cctcagatgataatccaaagatgagcctgagatgagacttccagaagacatgaaatg 4163  
4408 acagagataactcactcagagatgataatctgtcttccagatgataagagagatg 4467  
4164 acagagataactcactcagagatgataatctgtcttccagatgataagagagatg 4223  
4468 ctcaatcccaagatgagacatgagcagatgagactccttcttccagaatgagagatg 4527  
4224 ctcaatcccaagatgagacatgagcagatgagactccttcttccagaatgagagatg 4283  
4528 aaggttctcagagaaatcaagactcctcactcagatgagatgagacttccagacag 4587  
4284 aaggttctcagagaaatcaagactcctcactcagatgagatgagacttccagacag 4343  
4588 ttactgactgtcactcagatgagacatgagcagatgagacttccagatgagatgag 4647  
4344 ttactgactgtcactcagatgagacatgagcagatgagacttccagatgagatgag 4403  
4648 atgagatctcagagatgagcagacagagactcactcagatgagatgagatg 4699  
4404 atgagatctcagagatgagcagacagagactcactcagatgagatgagatg 4455

RESULT 4  
US-09-470-618-14  
: Sequence 14, Application US/09470618  
: Patent No. 6200560  
: GENERAL INFORMATION:  
: APPLICANT: Coulo, Linda B.  
: APPLICANT: Colosi, Peter C.  
: TITLE OF INVENTION: Adeno-Associated Vectors for Expression of Factor VIII  
: TITLE OF INVENTION: By Target Cells  
: FILE REFERENCE: Avigen-04082  
: CURRENT APPLICATION NUMBER: US/09/470,618  
: EARLIER FILING DATE: 1999-12-22  
: EARLIER APPLICATION NUMBER: 09/364,862  
: EARLIER FILING DATE: 1999-07-30  
: EARLIER APPLICATION NUMBER: 60/125,974  
: EARLIER FILING DATE: 1999-03-24  
: EARLIER APPLICATION NUMBER: 60/104,994  
: NUMBER OF SEQ ID NOS: 15  
: SOFTWARE: PatentIn Ver. 2.0  
: SEQ ID NO 14  
: LENGTH: 4999  
: TYPE: DNA  
: ORGANISM: Artificial Sequence



FEATURE:  
: OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
; Best Local Similarity 99.0%; DB 4; Length 4999;  
us-09-470-618-14 Pred. No. 0;  
Matches 4384; Conservative 0; Mismatches 2; Indels 42; Gaps 1;

```
Query Match      90.9%; Score 4330.8; DB 4; Length 4999;
Best Local Similarity 99.0%; DB 4; Length 4999;
Matches 4384; Conservative 0; Mismatches 2; Indels 42; Gaps 1;

QY 266 ccaccatgcacaatagagctctccacactgctctctctctgctctcttgcgattcgtctta 325
D 266 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
D 411 ccacatcgtcaaatagagctctccacactgctctctctctgctctcttgcgattcgtctta 470
QY 326 gtgcacacacagaagatctactctgtggtgagtggaactgtcatgtggaactatgtcaaatg 385
D 326 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
D 471 gtgcacacacagaagatctactctgtggtgagtggaactgtcatgtggaactatgtcaaatg 530
QY 386 atctcgtgagctgctgctgtgagcgaagaattcctcctctagagtgccaanaatcttccat 445
D 386 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
D 531 atctcgtgagctgctgctgtgagcgaagaattcctcctctagagtgccaanaatcttccat 590
QY 446 tcaaacactcagtcgtgtacaaaaaagactctgtttgtagaattcaagggttcaactttca 505
D 446 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
D 591 tcaaacactcagtcgtgtacaaaaaagactctgtttgtagaattcaagggttcaactttca 650
QY 506 acatcgcctaagccaaggccaccctggtggtgtcgtctaggtccctaccatccagctgag 565
D 506 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
D 651 acatcgcctaagccaaggccaccctggtggtgtcgtctaggtccctaccatccagctgag 710
QY 566 ttatgatcacagtgtgtcatctacttaagaacatggtctccatccctgtccagctctcatg 625
D 566 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
D 711 ttatgatcacagtgtgtcatctacttaagaacatggtctccatccctgtccagctctcatg 770
QY 626 ctgtgtgtgtatcctctactgtgaaagctctctgagggagctgaaatgatatgatacgaagc 685
D 626 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
D 771 ctgtgtgtgtatcctctactgtgaaagctctctgagggagctgaaatgatatgatacgaagc 830
QY 686 aagaaggaaagaagatgataaagttctccctgtgtggaagccatataatcttcggcagg 745
D 686 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
D 831 aagaaggaaagaagatgataaagttctccctgtgtggaagccatataatcttcggcagg 890
QY 746 tctctgaagaagatggttcccaatgtgcctctgacccactgtgcttactaccatactctt 805
D 746 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
D 891 tctctgaagaagatggttcccaatgtgcctctgacccactgtgcttactaccatactctt 950
QY 806 ctcaatgtgagactgttaaaagacttgaattcagagccctcatctgagccctactagatgta 865
D 806 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
D 951 ctcaatgtgagactgttaaaagacttgaattcagagccctcatctgagagccctactagatgta 1010
QY 866 gagaagaggaagctctgccaaggaagaacagacctgtgccaanaattatatactacttttg 925
D 866 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
D 1011 gagaagaggaagctctgccaaggaagaacagacctgtgccaanaattatatactacttttg 1070
QY 926 ctgtatctgtatgaaagggaagaagttgtgcactcagaacaanaaactccttgaatgcagata 985
D 926 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
D 1071 ctgtatctgtatgaaagggaagaagttgtgcactcagaacaanaaactccttgaatgcagata 1130
QY 986 ggaatgtgtcaatctgtcctgggctgtgccttaaatgtcacacagtcgaatgtgtatgtaaca 1045
D 986 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
D 1131 ggaatgtgtcaatctgtcctgggctgtgccttaaatgtcacacagtcgaatgtgtatgtaaca 1190
QY 1046 ggtctctgcacaggtctgtatgtgagtcgacagaagaatacagtcataatgtgcatgttgtaaa 1105
D 1046 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
D 1191 ggtctctgcacaggtctgtatgtgagtcgacagaagaatacagtcataatgtgcatgttgtaaa 1250
QY 1106 tgggacacacactcctgaagtggaactcaatatctcccgaaagtcacacacttctcttgagga 1165
D 1106 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
D 1251 tgggacacacactcctgaagtggaactcaatatctcccgaaagtcacacacttctcttgagga 1310
QY 1166 acacatgcacaggtctctctgtgaaatctgcgcaataaacttctcctaactgtcctcaaacactct 1225
D 1166 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
D 1311 acacatgcacaggtctctctgtgaaatctgcgcaataaacttctcctaactgtcctcaaacactct 1370
```

```
QY 1226 tgatggaaccttggaacagtttctactgtttgtcatatctctccaccacaatgatgca 1285
D 1226 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
D 1371 tgatggaaccttggaacagtttctactgtttgtcatatctctccaccacaatgatgca 1430
QY 1286 tgaagacttatgtcaaaagttagacagctgtgtccagagaagccccaactacgaatgaaata 1345
D 1286 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
D 1431 tgaagacttatgtcaaaagttagacagctgtgtccagagaagccccaactacgaatgaaata 1490
QY 1346 atgaaagaagcggaagactatgatgatctctactgtatcttcgaaatgtgatgtgcagt 1405
D 1346 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
D 1491 atgaaagaagcggaagactatgatgatctctactgtatcttcgaaatgtgatgtgtgcagt 1550
QY 1406 ttgatgtgacaactctctcctctcttatcccaattgcctcgaagttgtccaagaagatcccta 1465
D 1406 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
D 1551 ttgatgtgacaactctctcctctcttatcccaattgcctcgaagttgtccaagaagatcccta 1610
QY 1466 aaacttgggtacatctacatgtgtgtgaaagaaggaagctggaactatgtctccttaagtc 1525
D 1466 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
D 1611 aaacttgggtacatctacatgtgtgtgaaagaaggaagctggaactatgtctccttaagtc 1670
QY 1526 tgcgcccccgatgacagaagaatataaaagtcaatttgaacaatgtgccctcagcgatgtg 1585
D 1526 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
D 1671 tgcgcccccgatgacagaagaatataaaagtcaatttgaacaatgtgccctcagcgatgtg 1730
QY 1586 gtaggaagtatacaaaaagctcagatttatggtacatacagaagatgaaacctttagagctgtg 1645
D 1586 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
D 1731 gtaggaagtatacaaaaagctcagatttatggtacatacagaagatgaaacctttagagctgtg 1790
QY 1646 aagcttatcacaatgaaatcagaatctctggaacatttactttaggggaaagttagagaca 1705
D 1646 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
D 1791 aagcttatcacaatgaaatcagaatctctggaacatttactttaggggaaagttagagaca 1850
QY 1706 caactgtgatataatttaagaatacaagaagcagaacatatatacatctaccctcagaa 1765
D 1706 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
D 1851 caactgtgatataatttaagaatcaagaagcagaacatatatacatctaccctcagaa 1910
QY 1766 tcaactgtgtcgcgtcctctgtgtatcagaagagatataccaaaaggtgtaaaacttgaag 1825
D 1766 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
D 1911 tcaactgtgtcgcgtcctctgtgtatcagaagagatataccaaaaggtgtaaaacttgaag 1970
QY 1826 atttccaatctgtccagagaagaatatacaataataatgagacagtgactgtagaagatg 1885
D 1826 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
D 1971 atttccaatctgtccagagaagaatatacaataataatgagacagtgactgtagaagatg 2030
QY 1886 ggcacaactaaatcagatccctcgtgtgcctgacccgctatctactcagttcgttaatatg 1945
D 1886 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
D 2031 ggcacaactaaatcagatccctcgtgtgcctgacccgctatctactcagttcgttaatatg 2090
QY 1946 agagagatctagcttccaggaactatgtgcctctccatcatctgtacaaaagaatctgag 2005
D 1946 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
D 2091 agagagatctagcttccaggaactatgtgcctctccatcatctgtacaaaagaatctgag 2150
QY 2006 atcaaaagaggaagaacagaataatgtccaagaagaggaatgtcatcctgttctctatgt 2065
D 2006 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
D 2151 atcaaaagaggaagaacagaataatgtccaagaagaggaatgtcatcctgttctctatgt 2210
QY 2066 atgagaacccgaagctgtgtacccaagaagaataatacagcttcttccccaatccagctg 2125
D 2066 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
D 2211 atgagaacccgaagctgtgtacccaagaagaataatacagcttcttccccaatccagctg 2270
QY 2126 gagtgcagcttgagagatccgaaggttccaagcttccaacatcatgtcagaagcatgaatgct 2185
D 2126 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
D 2271 gagtgcagcttgagagatccgaaggttccaagcttccaacatcatgtcagaagcatgaatgct 2330
QY 2186 atgtttctgatagttttgcaatgtgtcagtttgttgtaataatgtgtgtaactgtgatactc 2245
D 2186 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
D 2331 atgtttctgatagttttgcaatgtgtcagtttgttgtaataatgtgtgtaactgtgatactc 2390
QY 2246 taagcatgtgagacacagactcctctctcgtctctctcctcctcgtgataactcctaacc 2305
D 2246 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
D 2391 taagcatgtgagacacagactcctctctcgtctctctcctcctcgtgataactcctaacc 2450
QY 2306 acaaaatgtctatgagaacacactcaccctatccatctcagagagaactgtctca 2365
```



Db 2451 acaaaatgctatgaaacacacactcaccccttcccaattccaggaagaattcttca 2510  
QY 2366 tctcgatgaaacccagctctatgatctctgggtgcccacactcagactctcgaaaca 2425  
Db 2511 tctcgatgaaacccagctctatgatctctgggtgcccacactcagactctcgaaaca 2570  
QY 2426 gaggcatgacccgcttactgaaaggttctcagtctgtcaacaagaacactggtaattacg 2485  
Db 2571 gaggcatgacccgcttactgaaaggttctcagtctgtcaacaagaacactggtaattacg 2630  
QY 2486 aggcacgtatgaaagatatcttcagcatactctgtcgaatcaaaaacatgcatctgaacca 2545  
Db 2631 aggcacgtatgaaagatatcttcagcatactctgtcgaatcaaaaacatgcatctgaacca 2690  
QY 2546 gaagcttcccaagattccaagacacccctagcactagcaaaagaatttaatgcccac 2605  
Db 2691 gaagcttcccaagattccaagacacccctagcactagcaaaagaatttaatgcccac 2708  
QY 2606 caccagctctgaaacgcacatcaacgcggaaataactcgtactactcttcagtcagatcaag 2665  
Db 2709 caccagctctgaaacgcacatcaacgcggaaataactcgtactactcttcagtcagatcaag 2768  
QY 2666 aggaatctgactatgatatataccatatacagttgaaatgaaaggaagaatttgacatt 2725  
Db 2769 aggaatctgactatgatatataccatatacagttgaaatgaaaggaagaatttgacatt 2828  
QY 2726 atgattgaaagtaaaatataagcccccgagcttccaagaagaacacacacactttaa 2785  
Db 2829 atgattgaaagtaaaatataagcccccgagcttccaagaagaacacacacactttaa 2888  
QY 2786 tctgctgcagctgagagagctctgagataatgagatgagtagctcccaactgttcaagaa 2845  
Db 2889 tctgctgcagctgagagagctctgagataatgagatgagtagctcccaactgttcaagaa 2948  
QY 2846 acaagagctcagaagctgagagctctcctcagttcaagaagaagttttccaaggaattacg 2905  
Db 2949 acaagagctcagaagctgagagctctcctcagttcaagaagaagttttccaaggaattacg 3008  
QY 2906 atgagctcttcaacacagccctataacgctgagaaactaaatgaacatttggaaactctg 2965  
Db 3009 atgagctcttcaacacagccctataacgctgagaaactaaatgaacatttggaaactctg 3068  
QY 2966 ggcacataataagacagaagcttgaaagataatacatggttaacttcagaaatcagacct 3025  
Db 3069 ggcacataataagacagaagcttgaaagataatacatggttaacttcagaaatcagacct 3128  
QY 3026 ctgcctcctatctcttctatctcagccttatcttctatgagaaagatcaagaagcaagag 3085  
Db 3129 ctgcctcctatctcttctatctcagccttatcttctatgagaaagatcaagaagcaagag 3188  
QY 3086 cagaacctaagaaaaaactctgccaagcctaatagaaccaaactactcttggaaagtgc 3145  
Db 3189 cagaacctaagaaaaaactctgccaagcctaatagaaccaaactactcttggaaagtgc 3248  
QY 3146 aacatacatggaacccaactaaagatgagcttgactgacaaagctctggcctatctctctg 3205  
Db 3249 aacatacatggaacccaactaaagatgagcttgactgacaaagctctggcctatctctctg 3308  
QY 3206 atgcttgaccttgaaaaaagatgacacagcctgattgagccctctcgtgctgcaca 3265  
Db 3309 atgcttgaccttgaaaaaagatgacacagcctgattgagccctctcgtgctgcaca 3368  
QY 3266 caaacacactgaacccctctcatgagagacaaagtacacagtacagaatttgcctgcttt 3325  
Db 3369 caaacacactgaacccctctcatgagagacaaagtacacagtacagaatttgcctgcttt 3428  
QY 3326 tcaacatcttctgataagaccaaagctggtactctcactgtaaaataatggaaagaaactgca 3385  
Db 3429 tcaacatcttctgataagaccaaagctggtactctcactgtaaaataatggaaagaaactgca 3488  
QY 3386 gggctccctgcaatatcacagatggaagatcccaactttaaagaagatatcgcttccatg 3445  
Db 3489 gggctccctgcaatatccagatggaagatcccaactttaaagaagatatcgcttccatg 3548  
QY 3446 caatcaatgctacataaataatgatacactacactgcttgatgataatgctcagagatcaagaa 3505  
Db 3549 caatcaatgctacataaataatgatacactacactgcttgatgataatgctcagagatcaagaa 3608  
QY 3506 ttcgattgatactgctcagacatgaggcagaatgaaaacatccatcttaattcactg 3565  
Db 3609 ttcgattgatactgctcagacatgaggcagaatgaaaacatccatcttaattcactg 3668  
QY 3566 gacatgctgttcaactgtaagaaaaaagaggatataaaatggacactgtacatctctatc 3625  
Db 3669 gacatgctgttcaactgtaagaaaaaagaggatataaaatggacactgtacatctctatc 3728  
QY 3626 caggtgttcttgagaacagtggaaatgttaccatcccaagcctggaatttggcgggtggat 3685  
Db 3729 caggtgttcttgagaacagtggaaatgttaccatcccaagcctggaatttggcgggtggat 3788  
QY 3686 gcttaattggcagacatcacaatgctgagatgagacacacttctcgtgtacagcaata 3745  
Db 3789 gcttaattggcagacatcacaatgctgagatgagacacacttctcgtgtacagcaata 3848  
QY 3746 agtctcagaatccctctgggaatggttcttgacacataagagatttccagattacagct 3805  
Db 3849 agtctcagaatccctctgggaatggttcttgacacataagagatttccagattacagct 3908  
QY 3806 caggaacatataagacagctgggcccacaagctggccagactcatatccagatcaata 3865  
Db 3909 caggaacatataagacagctgggcccacaagctggccagactcatatccagatcaata 3968  
QY 3866 atgcttgagacacaaagagaccccttctcttgatcaaggtgatactgttggaccaaata 3925  
Db 3969 atgcttgagacacaaagagaccccttctcttgatcaaggtgatactgttggaccaaata 4028  
QY 3926 ttattcacggacatcaagcccagggctgcccgctcaagaatttccagacctcacaatctc 3985  
Db 4029 ttattcacggacatcaagcccagggctgcccgctcaagaatttccagacctcacaatctc 4088  
QY 3986 agttatcatcatgataatgcttctgattggaagaagctggcagaactatccagagaaattcca 4045  
Db 4089 agttatcatcatgataatgcttctgattggaagaagctggcagaactatccagagaaattcca 4148  
QY 4046 ctggaaccttaatggtcttctcttgccaatgctgattcaatctcggatataaacacaatttc 4105  
Db 4149 ctggaaccttaatggtcttctcttgccaatgctgattcaatctcggatataaacacaatttc 4208  
QY 4106 ttaacctccaattatgctcgataacatccglttggacccaacacttatatagcatctgca 4165  
Db 4209 ttaacctccaattatgctcgataacatccglttggacccaacacttatatagcatctgca 4268  
QY 4166 gcaactctcgacatgagcttgatgggctgtgatttaaatagcttgacgataatgcatctggaa 4225  
Db 4269 gcaactctcgacatgagcttgatgggctgtgatttaaatagcttgacgataatgcatctggaa 4328  
QY 4226 tggagagtaaaagcaatatacagatgcaagatcaactgcttcaactacttaaccaatgct 4285  
Db 4329 tggagagtaaaagcaatatacagatgcaagatcaactgcttcaactacttaaccaatgct 4388  
QY 4286 ttggcaactggtctcctccaagaagcttgacttcaactccaaggaggagtaatgctcggga 4345  
Db 4388 ttggcaactggtctcctccaagaagcttgacttcaactccaaggaggagtaatgctcggga 4448  
QY 4346 gacctcaggtgataataacaaagagctgctgcaagtggaattccagaagaacatgaaag 4405  
Db 4449 gacctcaggtgataataacaaagagctgctgcaagtggaattccagaagaacatgaaag 4508  
QY 4406 tcaagagtaactacccaagagtaaaatctctgcttaccagacatgataatgaaagagct 4465  
Db 4509 tcaagagtaactacccaagagtaaaatctctgcttaccagacatgataatgaaagagct 4568  
QY 4466 tctcatctccagcagtaagaatgagccatcaatgagactctcttttccaagatggcaag 4525  
Db 4569 tctcatctccagcagtaagaatgagccatcaatgagactctcttttccaagatggcaag 4628



Oy	4526	taaaagtttttaagggaatacaagctcctttaacacctgtgttgaaactcttagaccac	4585
Db	4629	taaaagtttttaagggaatacaagctcctttaacacctgtgttgaaactcttagaccac	4688
Oy	4586	cgttcctgactcgtcaccttcgaattcacccccagaattgggtgcaccagatgacctga	4645
Db	4689	cgttcctgactcgtcaccttcgaattcacccccagaattgggtgcaccagatgacctga	4748
Oy	4646	ggatgtgaagttctgggctgcgagacagagacctactgactgcagc	4693
Db	4749	ggatgtgaagttctgggctgcgagacagagacctactgactgcagc	4796
RESULT	5		
US-09-364-862-14			
		: Sequence 14, Application US/09364862	
		: Patent No. 6221349	
		: GENERAL INFORMATION:	
		: APPLICANT: Couto, Linda B.	
		: APPLICANT: Colosi, Peter C.	
		: TITLE OF INVENTION: ADENO-ASSOCIATED VECTORS FOR EXPRESSION OF FACTOR VIII	
		: TITLE OF INVENTION: BY TARGET	
		: TITLE OF INVENTION: CELLS	
		: FILE REFERENCE: AVIGEN-03743	
		: CURRENT APPLICATION NUMBER: US/09/364,862	
		: CURRENT FILING DATE: 1999-07-30	
		: EARLIER APPLICATION NUMBER: 60/125,974	
		: EARLIER FILING DATE: 1999-03-24	
		: EARLIER APPLICATION NUMBER: 60/104,994	
		: EARLIER FILING DATE: 1998-10-20	
		: NUMBER OF SEQ ID NOS: 14	
		: SOFTWARE: PatentIn Ver. 2.0	
		: SEQ ID NO 14	
		: LENGTH: 4999	
		: TYPE: DNA	
		: ORGANISM: Artificial Sequence	
		: FEATURE:	
		: OTHER INFORMATION: Description of Artificial Sequence: Synthetic	
		: US-09-364-862-14	

Query Match	90.9%	Score 4330.8	DB 4	Length 4999
Best Local Similarity	99.0%	Pred. No. 0		
Matches 4384	Conservative 0	Mismatches 2	Indels 42	Gaps 1
QY 266	ccacatgcaataaataagctctccacctgctctctctctgctcttgcatctgctta	325		
Db 411	ccacatgcaataaataagctctccacctgctctctctctgctcttgcatctgctta	470		
QY 326	gtgcaccagaagaataactactcctgggtgcagtggaactgtcatlvggaactat	385		
Db 471	gtgcaccagaagaataactactcctgggtgcagtggaactgtcatlvggaactat	530		
QY 386	atctcggtgagctgctcctgtgacgcgaagaattcccccagaaatgccaatttttccat	445		
Db 531	atctcggtgagctgctcctgtgacgcgaagaattcccccagaaatgccaatttttccat	590		
QY 446	tcaaacacctcagtcgtgtfacaaaaaagaactcgtttgtgaattccaggttccatttca	505		
Db 591	tcaaacacctcagtcgtgtfacaaaaaagaactcgtttgtgaattccaggttccatttca	650		
QY 506	acatcgtcgaagccaagaccacccctgagtggtctgtctagtctccatccacagctgag	565		
Db 651	acatcgtcgaagccaagaccacccctgagtggtctgtctagtctccatccacagctgag	710		
QY 566	tttatgatacagtggtgtcaatcaactaaagaacatggtcccatccctgtcgaatttcagt	625		
Db 711	tttatgatacagtggtgtcaatcaactaaagaacatggtcccatccctgtcgaatttcagt	770		
QY 626	ctgtgtgtgtatcctactctggaagcctctctgaaggagctgaaatavagatcagaccagtc	685		
Db 771	ctgtgtgtgtatcctactctggaagcctctctgaaggagctgaaatavagatcagaccagtc	830		

OY	666	aaagvggaaagaagatagataaagctctccctcgtgtggaagccatacatatgtctgcgcag	745
Db	831	aaagvggaaagaagatagataaagctctccctcgtgtggaagccatacatatgtctgcgcag	890
OY	746	tcctcgaaagaagaatgctccaaatgagcccttgaccccaactgtgccttaactcatatctt	805
Db	891	tcctcgaaagaagaatgctccaaatgagcccttgaccccaactgtgccttaactcatatctt	950
OY	806	ctcatgtgagaccctgtgtaaaagacttgaattcaagcctcatctggagccctactagtatga	865
Db	951	ctcatgtgagaccctgtgtaaaagacttgaattcaagcctcatctggagccctactagtatga	1010
OY	866	gagaaagggaagctcgtccaaagaagaagaacagaccccttgacaacaatttaactactttt	925
Db	1011	gagaaagggaagctcgtccaaagaagaagaacagaccccttgacaacaatttaactactttt	1070
OY	926	ctgatattgataagagggaataaagatctgagacccagaataaagaagactctctgatagcagata	985
Db	1071	ctgatattgataagagggaataaagatctgagacccagaataaagaagactctctgatagcagata	1130
OY	966	gggagatgctgcatctcgtccgggccttgccctaaatgcaacacagtcacatggtatgtaaca	1045
Db	1131	gggagatgctgcatctcgtccgggccttgccctaaatgcaacacagtcacatggtatgtaaca	1190
OY	1046	ggtctctgcgcaggtcgtatattggaatgccaagagaaatccagatcatgtgagatgagaa	1105
Db	1191	ggtctctgcgcaggtcgtatattggaatgccaagagaaatccagatcatgtgagatgagaa	1250
OY	1106	tggtgcaaccaactcctgtaagctgacactcaaatctctcgaaagtcacaacattctcttgagga	1165
Db	1251	tggtgcaaccaactcctgtaagctgacactcaaatctctcgaaagtcacaacattctcttgagga	1310
OY	1166	accatctgcgaagcgtctcttggaatctcgcgaataaactcttcctactgctgaacaacct	1225
Db	1311	accatctgcgaagcgtctcttggaatctcgcgaataaactcttcctactgctgaacaacct	1370
OY	1226	tgatgtgacaccttggtgcagattcttctaactgtttgtgatactctctccacaacaatgaatgga	1285
Db	1371	tgatgtgacaccttggtgcagattcttctaactgtttgtgatactctctccacaacaatgaatgga	1430
OY	1286	tggtgaagcttatgtcaaaagttagaagaagctgtgcccgaaggaagaaacccaacttaagaatgaanaa	1345
Db	1431	tggtgaagcttatgtcaaaagttagaagaagctgtgcccgaaggaagaaacccaacttaagaatgaanaa	1490
OY	1346	atgaaagaagcggaaagactatgatagtatgatacttctgatactctgaatggaatggtgcagat	1405
Db	1491	atgaaagaagcggaaagactatgatagtatgatacttctgatactctgaatggaatggtgcagat	1550
OY	1406	ttgatgatgacaacctctctctctcttaaccaaatcgcgtcagttgccaagaagaactccta	1465
Db	1551	ttgatgatgacaacctctctctctcttaaccaaatcgcgtcagttgccaagaagaactccta	1610
OY	1466	aaactcttggtgaatcatatgaattgcgtcgtgaaagagagaaagactgggaactatgctcccttaagcc	1525
Db	1611	aaactcttggtgaatcatatgaattgcgtcgtgaaagagagaaagactgggaactatgctcccttaagcc	1670
OY	1526	tcgtcccccgatgaaagagaagtataaaagtcaaatatttgacaacatgagccctcagcggatgtg	1585
Db	1671	tcgtcccccgatgaaagagaagtataaaagtcaaatatttgacaacatgagccctcagcggatgtg	1730
OY	1586	gttaggaagtgacaataaagcttcggatttatgtgcataacacagatgaanaactttaagactcgtg	1645
Db	1731	gttaggaagtgacaataaagcttcggatttatgtgcataacacagatgaanaactttaagactcgtg	1790
OY	1646	aaagctatcatgacatgaaatccaggaatcttggtgaccccttaacttatgtgggaagtgtgagaca	1705
Db	1791	aaagctatcatgacatgaaatccaggaatcttggtgaccccttaacttatgtgggaagtgtgagaca	1850
OY	1706	caactgtgatataatttaagaatcaacagcaagcagacccatataaacttaacctcaacggaa	1765
Db	1851	caactgtgatataatttaagaatcaacagcaagcagacccatataaacttaacctcaacggaa	1910

QY	986	ggagtgctcgatctgctccgggctctggccttaaaatgcaacagtcgaatggttatgaaaca	1045
Db	1131	ggagatgcgtacatctgctccgggctctggccttaaaatgcaacagtcgaatggttatgaaaca	1190
QY	1046	ggtctctccgaaggtctgattgsgatgcaacaggaacacgaatcatatgagctgagattgga	1105
Db	1191	ggtctctccgaaggtctgattgsgatgcaacaggaacacgaatcatatgagctgagattgga	1250
QY	1106	tgggacacacatctctgaaagtcgaactcaatatctctcgaaagtcacaacattcttgtgaga	1165
Db	1251	tgggacacacatctctgaaagtcgaactcaatatctctcgaaagtcacaacattcttgtgaga	1310
QY	1166	acgatccgcagggcgctctcttggaaatctcgcaataacttctctactgtctcaaacact	1225
Db	1311	acgatccgcagggcgctctcttggaaatctcgcaataacttctctactgtctcaaacact	1370
QY	1226	tgatggaaccttgagacaattctctactgcttctgtcatatctctctcccaacaatgagtga	1285
Db	1371	tgatggaaccttgagacaagttctctactgcttctgtcatatctctctcccaacaatgagtga	1430
QY	1286	tggsaagcttatgtcaaaagtacagacagctgtctccagaggaaacccaactcgaatgaaaaa	1345
Db	1431	tggsaagcttatgtcaaaagtacagacagctgtctccagaggaaacccaactcgaatgaaaaa	1490
QY	1346	atggaagaagcgggaagactatgatagatacttactctgatacttctgaatgagtgtgcaggt	1405
Db	1491	atggaagaagcgggaagactatgatagatacttactctgatacttctgaatgagtgtgcaggt	1550
QY	1406	tgtgatagtacaactctctctctcttataccaattcgctcagttgcagaagactccta	1465
Db	1551	tgtgatagtacaactctctctctcttataccaattcgctcagttgcagaagactccta	1610
QY	1466	aaacttgygtaacattacatctgctcgtctgaaagagagactgagactatgctcccttagtcc	1525
Db	1611	aaacttgygtaacattacatctgctcgtctgaaagagagactgagactatgctcccttagtcc	1670
QY	1526	tcgcccccgatgacagagaagtataaaatgcataatttgaacaatgycctccagcgagattg	1585
Db	1671	tcgcccccgatgacagagaagtataaaatgcataatttgaacaatgycctccagcgagattg	1730
QY	1586	gtagaagaatcaaaaaagctccgatttatatgagcatacaacagatgaaaacttgaagactgtg	1645
Db	1731	gtagaagaatcaaaaaagctccgatttatatgagcatacaacagatgaaaacttgaagactgtg	1790
QY	1646	aagctattcagcatgaaatccagaatcttgygaccttactatttatggyggaagtltgagaca	1705
Db	1791	aagctattcagcatgaaatccagaatcttgygaccttactatttatggyggaagtltgagaca	1850
QY	1706	cacgtgtgatatattatgaataatcaagcaagcagacacatataaatctaacctcaacggaa	1765
Db	1851	cacgtgtgatatattatgaataatcaagcaagcagacacatataaatctaacctcaacggaa	1910



[illegible]

Db	2949	aacaggccagagtggcagtgccctcagttcaagaagtglttccaagatcttactg	3008
Qy	2906	atggctcccttaactcaagcccttaccgtyggaacataatgaacatttggagccctgg	2965
Db	3009	atggctcccttaactcaagcccttaccgtyggaacataatgaacatttggagccctgg	3068
Qy	2966	ggccalabataagacagaaagtgtgaagataataatcattglaacttcagaaatcagcct	3025
Db	3069	ggccalabataagacagaaagtgtgaagataataatcattglaacttcagaaatcagcct	3128
Qy	3026	ctggctccctatccctcttattctagccttattcttaataagaaatagaagagcagag	3085
Db	3129	ctggctccctatccctcttattctagccttattcttaataagaaatagaagagcagag	3188
Qy	3086	cagaacactagaanaaaactttgtccaagcctaataaaccanaacttacttttgaagatgc	3145
Db	3189	cagaacactagaanaaaactttgtccaagcctaataaaccanaacttacttttgaagatgc	3248
Qy	3146	aacatcalatgagccaccataagaatgagtgttgaactgcgaagccctggctattctctg	3205
Db	3249	aacatcalatgagccaccataagaatgagtgttgaactgcgaagccctggctattctctg	3308
Qy	3206	atgttgaccctggnaaaagatgtgcactccagcgttatgttgaccctctcgtctgcaca	3265
Db	3309	atgttgaccctggnaaaagatgtgcactccagcgttatgttgaccctctcgtctgcaca	3368
Qy	3266	ctaaacaactgaaccccgctccatcgggagagaatgtgacatacaggaaatttcctcgtttt	3325
Db	3369	ctaaacaactgaaccccgctccatcgggagagaatgtgacatacaggaaatttcctcgtttt	3428
Qy	3326	tcaccactcttbtgatgagaccacaagaagctggtactctcaactgtgaaataatggaagaactgcga	3385
Db	3429	tcaccactcttbtgatgagaccacaagaagctggtactctcaactgtgaaataatggaagaactgcga	3488
Qy	3386	gggctccctgcataatcccaatggaagatcccaactttaagaagaatatcgtcttcagt	3445
Db	3489	gggctccctgcataatcccaatggaagatcccaactttaagaagaatatcgtcttcagt	3548
Qy	3446	caatcaaatgtgtacataatgatgatacaactccctgtctatgattggtccaagatcaagaatga	3505
Db	3549	caatcaaatgtgtacataatgatgatacaactccctgtctatgattggtccaagatcaagaatga	3608
Qy	3506	ttcgatcgtatctcgtctcagcagatgaggcaatgaanaacatccatcttattcattcagt	3565
Db	3609	ttcgatcgtatctcgtctcagcagatgaggcaatgaanaacatccatcttattcattcagt	3668
Qy	3566	gaacttggttccactgttaccgnaaaaagagagatataaaatgagcactgttacaatctctatc	3625
Db	3669	gaacttggttccactgttaccgnaaaaagagagatataaaatgagcactgttacaatctctatc	3728
Qy	3626	cagaatgtttcttgagaacagtgtgaatgttccacatccaacatgaaatttggcgggtggat	3685
Db	3729	cagaatgtttcttgagaacagtgtgaatgttccacatccaacatgaaatttggcgggtggat	3788
Qy	3686	gaccttatttggcgaagcatctacatgcgcggagatgagcacactttctcgtgttgaagcaata	3745
Db	3789	gaccttatttggcgaagcatctacatgcgcggagatgagcacactttctcgtgttgaagcaata	3848
Qy	3746	agtgctcagaactccctcgtgggaatgagcttctcggacacatagagaatttcagattacagctt	3805
Db	3849	agtgctcagaactccctcgtgggaatgagcttctcggacacatagagaatttcagattacagctt	3908
Qy	3806	caggaaacaatbtgnaacagtgtggccccaagccttgccagactatcatctccggatccaatca	3865
Db	3909	caggaaacaatbtgnaacagtgtggccccaagccttgccagactatcatctccggatccaatca	3968
Qy	3866	atggctctgagacacaaagagaccccttctcttgatatacaagtgtgatatctgttggcaccaatga	3925
Db	3969	atggctctgagacacaaagagaccccttctcttgatatacaagtgtgatatctgttggcaccaatga	4028
Qy	3926	ttattctcggcatccaagaaccagaggtgtcccgctccaagaagtctccaagccttacaactctc	3985



Db	4029	ttattcaacgcatcaagaaccacaggtgtgccgctcagaagttcttcagcctctacatctctc	4088
Qy	3986	agttatcatcatgtaagtaagctctgtaatgtaggaaagaatgtagcagactatcatcgaagaattcca	4045
Db	4089	agttatcatcatgtaagtaagctctgtaagtaggaaagaatgtagcagactatcatcgaagaattcca	4148
Qy	4046	ctggaaccttaatggtctctctttggcaatgtagatcatctggaataaacacaaatctt	4105
Db	4149	ctggaaccttaatggtctctctttggcaatgtagatcatctggaataaacacaaatctt	4208
Qy	4106	ttaaccttcccaatattatgcttcggaataatccggtttgacccaactatattatgcttgcga	4165
Db	4209	ttaaccttcccaatattatgcttcggaataatccggtttgacccaactatattatgcttgcga	4268
Qy	4166	gcaactcttcgcataggaattgtagctgtaatttaaatgattgtagcgaatgtagcattggaa	4225
Db	4269	gcaactcttcgcataggaattgtagctgtaatttaaatgattgtagcgaatgtagcattggaa	4328
Qy	4226	tggagagtaagaacaatacatcagatgtagcagataactgtgtctatccctactttaaccaatagt	4285
Db	4329	tggagagtaagaacaatacatcagatgtagcagataactgtgtctatccctactttaaccaatagt	4388
Qy	4286	ttggaccactgtgtctcttcttaaaaagctcgcagcttccactccaaggagagtagatgctcgga	4345
Db	4389	ttggaccactgtgtctcttcttcttaaaaagctcgcagcttccactccaaggagagtagatgctcgga	4448
Qy	4346	gacctcagtggaataatccaaaagaatgtagctgtaagtgacttccagaagaacaatgaaag	4405
Db	4449	gacctcagtggaataatccaaaagaatgtagctgtaagtgacttccagaagaacaatgaaag	4508
Qy	4406	tcacagaagtaactactctcagggaggtaaatatctctgcttaccagaacaatgatatgtgaagaagt	4465
Db	4509	tcacagaagtaactactctcagggaggtaaatatctctgcttaccagaacaatgatatgtgaagaagt	4568
Qy	4466	ttcccatctctccagaatccaagaatgtggccatcagtggaactctctttttccaagaatgacaag	4525
Db	4569	ttcccatctctccagaatccaagaatgtggccatcagtggaactctctttttccaagaatgacaag	4628
Qy	4526	taaaagtcttccagggaataatcaagaactctcttaccacccctgtgtgaactctcttagaccac	4585
Db	4629	taaaagtcttccagggaataatcaagaactctcttaccacccctgtgtgaactctcttagaccac	4688
Qy	4586	cgttactgactgcgtacactctggaattcaatccacccacagaattggtgacacagaatgtagccttga	4645
Db	4689	cgttactgactgcgtacactctggaattcaatccacccacagaattggtgacacagaatgtagccttga	4748
Qy	4646	ggataggaaggtctcggcgtcgcagggcacaaggaactctactgactgcagac	4693
Db	4749	ggataggaaggtctcggcgtcgcagggcacaaggaactctactgactgcagac	4796
RESULT 6			
US-09-470-618-13			
: Sequence 13, Application US/09470618			
: Patent No. 6200560			
: GENERAL INFORMATION:			
: APPLICANT: Colosi, Linda B.			
: APPLICANT: Colosi, Peter C.			
: TITLE OF INVENTION: Adeno-Associated Vectors for Expression of Factor VIII			
: FILE REFERENCE: Avigen-04082			
: CURRENT APPLICATION NUMBER: US/09/470, 618			
: CURRENT FILING DATE: 1999-12-22			
: EARLIER APPLICATION NUMBER: 09/364, 862			
: EARLIER FILING DATE: 1999-07-30			
: EARLIER APPLICATION NUMBER: 60/125, 974			
: EARLIER FILING DATE: 1999-03-24			
: EARLIER APPLICATION NUMBER: 60/104, 994			
: EARLIER FILING DATE: 1998-10-20			
: NUMBER OF SEQ. ID NOS: 15			
: SOFTWARE: Patentlin Ver. 2.0			
: SEQ ID NO 13			
: LENGTH: 11933			

[illegible]



Db	1322	accatctgcgaagcgcttcctcttggaatctccgcgaataacttcttcactgctcctaaacactct	1301
Qy	1226	tcgatggaccttggacagcttctctaactgtttctgtcatactctctccaccacaatgatggca	1285
Db	1382	tgaatggaccttggacagcttctctaactgtttctgtcatactctctccaccacaatgatggca	1441
Qy	1286	ttggaagcttatgtccaaagtatgaacagctgtctccaggggaaccccaactcagaaatgaanaata	1345
Db	1442	tggaaagcttatgtccaaagtatgaacagctgtctccaggggaaccccaactcagaaatgaanaata	1501
Qy	1346	atggaagaagcggagagactatgatgatgactcttaactgtatctcggaaatgatgttgctcaggt	1405
Db	1502	atggaagaagcggagagactatgatgatgactcttaactgtatctcggaaatgatgttgctcaggt	1561
Qy	1406	ctgatgatgacaacactctcctctctcttataccaactctgctcagttgtccagaagaagatcccta	1465
Db	1562	ctgatgatgacaacactctcctctctcttataccaactctgctcagttgtccagaagaagatcccta	1621
Qy	1466	aaacttggttaccttaccattgtctgccttgaagaaggagactggtgactatgcttcctctagtcc	1525
Db	1622	aaacttggttaccttaccattgtctgccttgaagaaggagactggtgactatgcttcctctagtcc	1681
Qy	1526	tcgcccccgatgacagaagtatataaagtatcaatttggacaacatggtccctcagcggagt	1585
Db	1682	tcgcccccgatgacagaagtatataaagtatcaatttggacaacatggtccctcagcggagt	1741
Qy	1586	gttagaagtatacaaaaagtcgcgaattta.tggtcatcacagata.tgaaacctttaaagctcgtg	1645
Db	1742	gttagaagtatacaaaaagtcgcgaattta.tggtcatcacagata.tgaaacctttaaagctcgtg	1801
Qy	1646	aagctatctcaagcatatgaatcagaatctcttggaccttaacttata.tggtgaagtgttgagaca	1705
Db	1802	aagctatctcaagcatatgaatcagaatctcttggaccttaacttata.tggtgaagtgttgagaca	1861
Qy	1706	caactgttgattatatttaagaataccaagaacagacacataataatccactcaaggaa	1765
Db	1862	caactgttgattatatttaagaataccaagaacagacacataataatccactcaaggaa	1921
Qy	1766	tcactgatgtccgtccctctgtcatctcaagggattataccaagaagtgtataaactcttgaaag	1825
Db	1922	tcactgatgtccgtccctctgtcatctcaagggattataccaagaagtgtataaactcttgaaag	1981
Qy	1826	attttccaactctgtccagagagaataatccaataataa.tggtcagatgtatgtatagaatgt	1885
Db	1982	attttccaactctgtccagagagaataatccaataataa.tggtcagatgtatgtatagaatgt	2041
Qy	1886	ggccaactaataatcagaatctcccggtgtcccttgaacccgctatactcctcagtttctgttaatgt	1945
Db	2042	ggccaactaataatcagaatctcccggtgtcccttgaacccgctatactcctcagtttctgttaatgt	2101
Qy	1946	agaagagatctagaatcttcaagagatcaat.tggtccctctcctcactgtcttaacaagaatctgtag	2005
Db	2102	agaagagatctagaatcttcaagagatcaat.tggtccctctcctcactgtcttaacaagaatctgtag	2161
Qy	2006	atccaaagatggaaaaccagataatgtctcagaacaagggaatgtcaatccgtttctctgtattgt	2065
Db	2162	atccaaagatggaaaaccagataatgtctcagaacaagggaatgtcaatccgtttctctgtattgt	2221
Qy	2066	atggaagaacggaaagctgtgtacactcaacagagaata.taaacgcttctctcccaatccagctg	2125
Db	2222	atggaagaacggaaagctgtgtacactcaacagagaata.taaacgcttctctcccaatccagctg	2281
Qy	2126	gagtcagactctgagagatccagagatcccaaacctccaacatcatatgacagcatcaatgtgt	2185
Db	2282	gagtcagactctgagagatccagagatcccaaacctccaacatcatatgacagcatcaatgtgt	2341
Qy	2186	atgttttctgatagtttctgcagttgtctcagtttctgtcatatgaggtgtgcatatctgtatactc	2245
Db	2342	atgttttctgatagtttctgcagttgtctcagtttctgtcatatgaggtgtgcatatctgtatactc	2401
Qy	2246	taagcatctggagacacagactgtaactcctctctgtctctctctctctgtatataactcttaaac	2305
Db	2402	taagcatctggagacacagactgtaactcctctctgtctctctctctctgtatataactcttaaac	2461

[illegible]







|||||  
Db 782 ctgtgtgtatctctactcgtgaagctctctgaggcgatgatgatcatcagacagctc 841  
QY 686 aaaaagggaagaagaagtataaagctctccctggtggaagccaatacatgtctgcagg 745  
Db 842 aaaaagggaagaagaagtataaagctctccctggtggaagccaatacatgtctgcagg 901  
QY 746 tccggaagaagaagtatcccaatggtcctctgacccactgtgccttcaactcatactct 805  
Db 902 tccggaagaagaagtatcccaatggtcctctgacccactgtgccttcaactcatactct 961  
QY 806 ctcatctggaacctgtgtaaaagacttgaatccaagcctcatctggagccctactagtatgta 865  
Db 962 ctcatctggaacctgtgtaaaagacttgaatccaagcctcatctggagccctactagtatgta 1021  
QY 866 gagaagggaagctctgccaagaagaagaagacagaccttgcaacaatttaactacttttg 925  
Db 1022 gagaagggaagctctgccaagaagaagaagacagaccttgcaacaatttaactacttttg 1081  
QY 926 ctgtattctgtaagggaagaagcttggtcactcagaacaagaagactccttgatgcagata 985  
Db 1082 ctgtattctgtaagggaagaagcttggtcactcagaacaagaagactccttgatgcagata 1141  
QY 986 ggagctgtgcatctgtctgggacctggtcctaaatgacacagtcatacgtgtatgtaaca 1045  
Db 1142 ggagctgtgcatctgtctgggacctggtcctaaatgacacagtcatacgtgtatgtaaca 1201  
QY 1046 ggtctctgcgcaggtctgtatgtaatgcacagaagaatcagctatgtggaatggttgtaa 1105  
Db 1202 ggtctctgcgcaggtctgtatgtaatgcacagaagaatcagctatgtggaatggttgtaa 1261  
QY 1106 tgggacacacactccgtgaagtgtgactcaataatctctcgaaggtcacacacattctgtgaa 1165  
Db 1262 tgggacacacactccgtgaagtgtgactcaataatctctcgaaggtcacacacattctgtgaa 1321  
QY 1166 acctctgcgcaggtctctgtgaaatctgcgaataacttctctactgtctcaaacactct 1225  
Db 1322 acctctgcgcaggtctctgtgaaatctgcgaataacttctctactgtctcaaacactct 1381  
QY 1226 tgaatgagacctggaagcttctactgtcttgcatactctctcccaacaatactgtgca 1285  
Db 1382 tgaatgagacctggaagcttctactgtcttgcatactctctcccaacaatactgtgca 1441  
QY 1286 tgggaagcttatgtcaaaagttagacagctgtgccagaggaaccccaactacgaatgtaaaaa 1345  
Db 1442 tgggaagcttatgtcaaaagttagacagctgtgccagaggaaccccaactacgaatgtaaaaa 1501  
QY 1346 atggaagaagcgggaagactatgatacttactgtatctcgtgaatgtgatagtgtgcagct 1405  
Db 1502 atggaagaagcgggaagactatgatacttactgtatctcgtgaatgtgatagtgtgcagct 1561  
QY 1406 ttgtatgtgaacaactctctctcccttatccaaatcgcgtcagttgccaagaagcatcccta 1465  
Db 1562 ttgtatgtgaacaactctctctcccttatccaaatcgcgtcagttgccaagaagcatcccta 1621  
QY 1466 aaacttggtaacatatactgtctgtaagaaggagagacttgggactatgtcctcctagtcc 1525  
Db 1622 aaacttggtaacatatactgtctgtaagaaggagagacttgggactatgtcctcctagtcc 1681  
QY 1526 tcgcccccgatgacagaagttataaagtcataattgaaacaatgccccctcagcgagctg 1585  
Db 1682 tcgcccccgatgacagaagttataaagtcataattgaaacaatgccccctcagcgagctg 1741  
QY 1586 gtggaagaagtaaaaaaagctcgattatgtgcatacacagatgaaaaccttlaagaactgtg 1645  
Db 1742 gtggaagaagtaaaaaaagctcgattatgtgcatacacagatgaaaaccttlaagaactgtg 1801  
QY 1646 aagctatcagcatgtaatcaggaatctctggagaccttacttatagtgggaagtgtgagaca 1705  
Db 1802 aagctatcagcatgtaatcaggaatctctggagaccttacttatagtgggaagtgtgagaca 1861  
QY 1706 cactgtgtattatatcttaagaatacgaagcagacataaacttactaccttcacgaa 1765  
|||||

Db 1862 cactgtgtattatatcttaagaatacgaagcagacacataaactatcaccttcacgaa 1921  
QY 1766 tcaactgagctccgtctctgttatccaagaagattacaacaaaggtgtataaacattggaag 1825  
Db 1922 tcaactgagctccgtctctgttatccaagaagattacaacaaaggtgtataaacattggaag 1981  
QY 1826 atttccaattctgcaggaagaatatcatcaataataatgacagctgactgtagaagatg 1885  
Db 1982 atttccaattctgcaggaagaatatcatcaataataatgacagctgactgtagaagatg 2041  
QY 1886 ggcacaactaaatcagatctctgtgtgcagaccgctatctactagtgttctgtaatttg 1945  
Db 2042 ggcacaactaaatcagatctctgtgtgcagaccgctatctactagtgttctgtaatttg 2101  
QY 1946 agagatctagcttcagagcactatgtgcctcctccatcatctgtctacaacaaactgtag 2005  
Db 2102 agagatctagcttcagagcactatgtgcctcctccatcatctgtctacaacaaactgtag 2161  
QY 2006 atcaagaaggaaacagataatgltcagaacaaggaaatgtcatctgttctgtatttg 2065  
Db 2162 atcaagaaggaaacagataatgltcagaacaaggaaatgtcatctgttctgtatttg 2221  
QY 2066 atgaaacccgaagctgtgtactctacaagaataatacaagcttctcccaatccagctg 2125  
Db 2222 atgaaacccgaagctgtgtactctacaagaataatacaagcttctcccaatccagctg 2281  
QY 2126 gagtgcagcttggaagatccagatctcagatcccaagctcccaatcatgtcacagcatatgct 2185  
Db 2282 gagtgcagcttggaagatccagatctcagatcccaagctcccaatcatgtcacagcatatgct 2341  
QY 2186 atgttttgataagtttgagttgtaagtttgatgtaagtggtgcaactgtgtaatc 2245  
Db 2342 atgttttgataagtttgagttgtaagtttgatgtaagtggtgcaactgtgtaatc 2401  
QY 2246 taagatctggagacagactgactctcttctgtctctctctctctctctctctctctctctct 2305  
Db 2402 taagatctggagacagactgactctcttctgtctctctctctctctctctctctctctctct 2461  
QY 2306 acaaaatgctctatgaagaacacacactacactatccacttccacttcagaggaactgtctca 2365  
Db 2462 acaaaatgctctatgaagaacacacactacactatccacttccacttccaggaagactgtctca 2521  
QY 2366 tgcgataggaaaacccaggtctatagatctctgggtgtcccaactcagacttctcgaaaca 2425  
Db 2522 tgcgataggaaaacccaggtctatagatctctgggtgtcccaactcagacttctcgaaaca 2581  
QY 2426 gaggcataccgcttacttgaaggttcttagtttgtaaaagaacactgtgtattatgag 2485  
Db 2582 gaggcataccgcttacttgaaggttcttagtttgtaaaagaacactgtgtattatgag 2641  
QY 2486 aggaagttatgaagataatccaacatactgtcgtgagttaaaaaaatgccaattgaaacaa 2545  
Db 2642 aggaagttatgaagataatccaacatactgtcgtgagttaaaaaaatgccaattgaaacaa 2701  
QY 2546 gaagcttctccagaatccaagacacccctagcactaggcaaaagcaatttaatgccaacc 2605  
Db 2702 gaagcttctccagaatccaagacacccctagcactaggcaaaagcaatttaatgccaacc 2709  
QY 2606 caccagctcttgaaaacgcatacaggggaataactgtactactcttcagtcagatgaag 2665  
Db 2710 -----gaaataactgttactactcttccagtcagatgaag 2743  
QY 2666 aggaatctgactatgatataccatatcagttgtgaatgaaagaagaattttacattc 2725  
Db 2744 aggaatctgactatgatataccatatcagttgtgaatgaaagaagaattttgacattc 2803  
QY 2726 atgatgagatgtaaaatcagaagcccccagacttccaagaagaacacgcacataattta 2785  
Db 2804 atgatgagatgtaaaatcagaagcccccagacttccaagaagaacacgcacataattta 2863  
QY 2786 ttgctgcagtggaagagctctgggattatgtagagatgtagtccccaatggttctaagaa 2845  
Db 2864 ttgctgcagtggaagagctctgggattatgtagagatgtagtccccaatggttctaagaa 2923  
|||||



2846 acagggctacagatggcagctgtccctcagttccaagaagattgtttccagaatttactg 2905  
|||||  
2924 acagggctacagatggcagctgtccctcagttccaagaagattgtttccagaatttactg 2963  
2906 atggctccttactcagcccttataccgttggagaactaaatgaacatttggactcttg 2965  
|||||  
2984 atggctccttactcagcccttataccgttggagaactaaatgaacatttggactcttg 3043  
2966 ggcacatataagaagcagaagtgtgaagataatcatctgttaactttcagaataaggcct 3025  
|||||  
3044 ggcacatataagaagcagaagtgtgaagataatcatctgttaactttcagaataaggcct 3103  
3026 ctgctccctattccttcttattcttagccttatttttttttttttttttttttttttttttt 3085  
|||||  
3104 ctgctccctattccttcttattcttagccttatttttttttttttttttttttttttttttt 3163  
3086 cagaacctagaanaaaactttgtcagaacctaaatgaacccaacttacttttggaaagtg 3145  
|||||  
3164 cagaacctagaanaaaactttgtcagaacctaaatgaacccaacttacttttggaaagtg 3223  
3146 aacatataatggcaccacaaatgaatgttgaactgaagcctgggtcttattctctg 3205  
|||||  
3224 aacatataatggcaccacaaatgaatgttgaactgaagcctgggtcttattctctg 3283  
3206 atgttgacctggaaaagaatgtgactcagggcctgatttggaccctctgtctgcaca 3265  
|||||  
3284 atgttgacctggaaaagaatgtgactcagggcctgatttggaccctctgtctgcaca 3343  
3266 ctacaacacttgaacctgtctcatgtggagacaaatgtacagatcacagaaatttgcctgttt 3325  
|||||  
3344 ctacaacacttgaacctgtctcatgtggagacaaatgtacagatcacagaaatttgcctgttt 3403  
3326 tccacatcttcttgaatgaacccaagctgtgacttccactgtaataatgtgaagaactctgca 3385  
|||||  
3404 tccacatcttcttgaatgaacccaagctgtgacttccactgtaataatgtgaagaactctgca 3463  
3386 gggctccctgaatcagatgaagaatccacattttaagaagaattatccgtcttcattg 3445  
|||||  
3464 gggctccctgaatcagatgaagaatccacattttaagaagaattatccgtcttcattg 3523  
3446 caatcaatggctacataatgaatgaatcactcgttgaatgaatggcagaatcaagaaga 3505  
|||||  
3524 caatcaatggctacataatgaatgaatcactcgttgaatgaatggcagaatcaagaaga 3583  
3506 ttgcatgtatctgtctcagatgaatggcagaatgaacacatccatcttcatcttcagt 3565  
|||||  
3584 ttgcatgtatctgtctcagatgaatggcagaatgaacacatccatcttcatcttcagt 3643  
3566 gacatgttctcagatgaatgaagaagaatgaatgaatgaatgaatgaatgaatgaatgaatga 3625  
|||||  
3644 gacatgttctcagatgaatgaagaagaatgaatgaatgaatgaatgaatgaatgaatgaatga 3703  
3626 cagggttttttggagacagttggaatgttaccatccaaagcttggaaatttggcgggtggaat 3685  
|||||  
3704 cagggttttttggagacagttggaatgttaccatccaaagcttggaaatttggcgggtggaat 3763  
3686 ggccttatctggcggacatcatctgtctggatgagcacacttttctgtgttaacagcata 3745  
|||||  
3764 ggccttatctggcggacatcatctgtctggatgagcacacttttctgtgttaacagcata 3823  
3746 agtttcagactcccttggagatgagcttggacacatgaagatttccaaatttcaacatt 3805  
|||||  
3824 agtttcagactcccttggagatgagcttggacacatgaagatttccaaatttcaacatt 3883  
3806 caggacacatctggacagttggccccaagcttggccagacttcatcttccggatcaatca 3865  
|||||  
3884 caggacacatctggacagttggccccaagcttggccagacttcatcttccggatcaatca 3943  
3866 atgctctggagacacaaaggaaccccttctctgtgataaggtgagatctgttggcaccaatga 3925  
|||||  
3944 atgctctggagacacaaaggaaccccttctctgtgataaggtgagatctgttggcaccaatga 4003

QY 3926 ttattcagcgcatcaagcaccagggtgccgtcagaagattctccagccttactcttc 3985  
|||||  
Db 4004 ttattcagcgcatcaagcaccagggtgccgtcagaagattctccagccttactcttc 4063  
3986 agtttaccatcatgttatagcttcttgaatgaagaagtggcagacttaccggaatttcca 4045  
|||||  
Db 4064 agtttaccatcatgttatagcttcttgaatgaagaagtggcagacttaccggaatttcca 4123  
4046 ctggaaccttaatgttcttcttcttggcaatgttgatctcatctgggataaacaacattt 4105  
|||||  
Db 4124 ctggaaccttaatgttcttcttcttggcaatgttgatctcatctgggataaacaacattt 4183  
4106 tttaacctccaatattgtcttcgcatatccgttttgacccaactccatattagccttgcga 4165  
|||||  
Db 4184 tttaacctccaatattgtcttcgcatatccgttttgacccaactccatattagccttgcga 4243  
4166 gcaactctcgcattggagttgatggcctgtgatttaaaatgttgcagatgcatttggaa 4225  
|||||  
Db 4244 gcaactctcgcattggagttgatggcctgtgatttaaaatgttgcagatgcatttggaa 4303  
4226 tggagagtaagaacatatcagatgcagatctactgttcaacttaccatattgt 4285  
|||||  
Db 4304 tggagagtaagaacatatcagatgcagatctactgttcaacttaccatattgt 4363  
4286 ttggcaccctgtctccttccaagaactcgaacttccaccgaaggagatgaatgcctgga 4345  
|||||  
Db 4364 ttggcaccctgtctccttccaagaactcgaacttccaccgaaggagatgaatgcctgga 4423  
4346 gacatcagttgaataatccaaaagctgtgctcagaatggacttccaaagaacatgaag 4405  
|||||  
Db 4424 gacatcagttgaataatccaaaagctgtgctcagaatggacttccaaagaacatgaag 4483  
4406 tcaagaagatgaactaccacaaagctgtgctcagaatggacttccaaagaacatgaag 4465  
|||||  
Db 4484 tcaagaagatgaactaccacaaagctgtgctcagaatggacttccaaagaacatgaag 4543  
4466 tcccatctccagcagatgaatggcagatgagatgagcttcttttcagaatggcaag 4525  
|||||  
Db 4544 tcccatctccagcagatgaatggcagatgagatgagatgagcttcttttcagaatggcaag 4603  
4526 taaaggttttccaggggaatacgaactccttcaacacgttgttgaactcttagaccac 4585  
|||||  
Db 4604 taaaggttttccaggggaatacgaactccttcaacacgttgttgaactcttagaccac 4663  
4586 cgttactgactcgtctactcgaatcaccacagagttgggtgacagatgtgcctga 4645  
|||||  
Db 4664 cgttactgactcgtctactcgaatcaccacagagttgggtgacagatgtgcctga 4723  
4646 ggaatggaggttctgtggttggcagacaggaaccttacttactgcag 4692  
|||||  
Db 4724 ggaatggaggttctgtggttggcagacaggaaccttacttactgcag 4770

RESULT 8  
US-08-882-083-1  
; Sequence 1, Application US/08882083  
; Patent No. 5869292  
; GENERAL INFORMATION:  
; APPLICANT: VOORBERG, Johannes J.  
; TITLE OF INVENTION: HYBRID PROTEINS WITH MODIFIED ACTIVITY  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30



```

1      CURRENT APPLICATION DATA:
2      APPLICATION NUMBER:  US/08/882,083
3      FILING DATE:
4      CLASSIFICATION:  514
5      PRIOR APPLICATION DATA:
6      APPLICATION NUMBER:  US 08/558,107
7      FILING DATE:  13-NOV-1995
8      ATTORNEY/AGENT INFORMATION:
9      NAME:  ISACSON, John P.
10     REGISTRATION NUMBER:  33,715
11     REFERENCE/DOCKET NUMBER:  30472/212
12     TELECOMMUNICATION INFORMATION:
13     TELEPHONE:  (202)672-5300
14     TELEFAX:  (202)672-5399
15     TELEX:  904136
16     INFORMATION FOR SEQ ID NO:  1:
17     SEQUENCE CHARACTERISTICS:
18     LENGTH:  5035 base pairs
19     TYPE:  nucleic acid
20     STRANDEDNESS:  single
21     TOPOLOGY:  linear
22     FEATURE:
23     NAME/KEY:  CDS
24     LOCATION:  35..5017
25     US-08-882-083-1

```

Query Match	78.78;	Score 3748;	DB 2;	Length 5035;
Best Local Similarity	87.48;	Pred. No. 0;		
Matches 4374; Conservative	0;	Mismatches 60;	Indels 570;	Gaps 2

OY	266	caaatgtgaataatagagctctcaacctgtgtctctctctgtgacctttgtgaccttgcgtta	325
Db	30	CCACCATTGGAAATAGAGCTCTCCACCTGCTTCTTTGTGTGCCATTTCTGCTTAA	89
OY	326	gtgcacaccagaagatactactccttggtgtgcagtgcgaactgtcataatgtgacagtg	385
Db	90	GTGCCACCGAAGATACCTACTGCGGTGGACAGTGGAACTGTCAATGGGACTATATGCCAAAGG	149
OY	386	atctcgggtgaagctgcctctgtgagcgcaagaattccctctcagatgtgcgaataactcttcac	445
Db	150	ATCTCGGTGAGCTGCTCTGTGGACGCAAGATTTCCTCTCTAGAGTGCCTTTTCCAT	209
OY	446	tcaacaacctcagtcgltgtcaaaaagaactcgtlgttagaatcacggttcaacctltta	505
Db	210	TCAACAACCTCACTCGGTGTCACAAAAGACCTGTTGTAAATTCACAGGTATCACCTTTTCA	269
OY	506	acatcgtcgaagcgaagcgcaacctgtgaatgggtctgtctagtgctctacatccagagctgaag	565
Db	270	ACATCGCTTAGCCAAAGGCCACCTGGATGGGTGCTGTAGTCTCTACCATCCAGGCTGAAG	329
OY	566	tttatgatcaagtgyltalttacactaaagaacatggtctccatccctgtcagtlcttcag	625
Db	330	TTTATGATACATGTGGTCATTACCTTAAGAACATGGCTTCCACTCTGTACAGCTTCATG	389
OY	626	ctgttggtgtatacctactgtgaaagctcttcgagggagctgaataatgtgtcagaccagtc	685
Db	390	CTGTTGTGTATCTACTACTGTGAAGCTTCTGAGGAGCTTAATATGATGATCAGACCACTG	449
OY	686	aaaggagagaagaatgtgataaagctctccctctgtgtgagagccatatatgtctctgcaag	745
Db	450	AAAGGAGAAAGAAGATGATTAAGTCTTCCCTGCTGTGGAAGCCATACATATGTCGTGGCAG	509
OY	746	tcctgtaaagaagaatgtgtccaatgtgcctctgcagccactgtgctcttacctatactctt	805
Db	510	TCTCTAATAAGATGGTCCAAATGTGGCTCTGAGCCACTGTGCTTACTCTACTATATCTTT	569
OY	806	ctcagctgtgacctgtgtaaaagaacttgaattcgaagcctctctgtgaagccctacagtatgt	865
Db	570	CTCATGTGAGACTGTGTAAAGACTTGAATTTCAGGCTCTCATTTGAGGCCCTACTAGTATGT	629
OY	866	gaggaaggagctctgccaaggaagaacagacctgtcaacaattatactacttlttg	925

Dp	630	GAGAAAGGAGTGTGGCCAAAGSAAAAACACAGACCTTGCACAAATTTATATACACTTTTTG	689
Qy	926	ctgtatcttgatgaagggaaaagtttgcacttcagaacaagaaccccttgatgcaagata	985
Dp	690	CTGTATTTTGATGAAGGGAAAAAGTTGGCACTCAGAACAAAAAGAACTCTTGATGCAAGATA	749
Qy	986	ggagatgcgcacttcgctcggggcccgccctaaatgcaacagtcgaatggtttctgtaaca	1045
Dp	750	GGGATGCTGCACTTCGCTCGGGCCCTGGCCCTTAATATGCAACACAGTCAATTTGTTATGAACA	809
Qy	1046	ggctctcgcgcagtgcttgattgaaatgcaacaggaatcaagtcctcttgcaatgtagtggaa	1105
Dp	810	GGTCTCTGCGCAAGTGTGATTTGGATGGCCACAGAAATAGTCTTTGGCATGTGATTTGGAA	869
Qy	1106	tgggcaacacatccctgaagtgcactcaatctcctcgaaagtgcaacatctcttgtagga	1165
Dp	870	TGGGCAACCACTCTGGAATGCACTCAATTTCCGAAAGGTGCACATTTCTTTGTGAGAGA	929
Qy	1166	accatcgcccaagcgctcccttggaatctcgcaataacttccctactgctgaacaacct	1222
Dp	930	ACCATCGCCAGGCGCTCTGGAAATTCGGCCAAATATCTTTCCTTACTGCTCAACACACT	989
Qy	1226	tgaatgaccttggaaagttctactgtcttgcaatctcttcacacaacatgtagga	1288
Dp	990	TGATGGACCTTGAGACAGTTTACTGTGTTTGTCAATCTTCTCCACCAACATGATGGCA	1049
Qy	1286	tggaaagcttatgtcaaaagtagacagctgtgcacaaggaaacccaactacgaatgtaaaaaa	1345
Dp	1050	TGGAGCTTATGTCAAGTAGACAGTGTCCAGAGAACCCCACTACGAATGAAAAATA	1109
Qy	1346	atgaaagaacggaagactatgaatgaatctctcgtatctcgaatctgatatgtgtcgaag	1405
Dp	1110	ATGAAGAAGCGGAAGACTATGATGATGATCTTACTGATTTCTGAAAGATGTGTGACAGT	1169
Qy	1406	ttagatgataaacactccctccctcttcaacaatcgctcaagttgccaagaagactccta	1465
Dp	1170	TGTGATGATGACAACTCTCTCTTTCCTTTATCCAAATGGCTCAGTGGCCAAAGCAATCCTA	1229
Qy	1466	aaacttggatcatcatcatctgctgtcgaagaaggagacttggatcatgtctccctagctcc	1525
Dp	1230	AAACTTGGGTACATTTACTATCTCTCTGTAAGAGGAGCATGGAGCATGTGCTCCCTTAGTCC	1288
Qy	1526	tcgcccccgatgacagaagaagtataaaagtcataatttgaaacaatgagccctcagcgatg	1588
Dp	1290	TGCGCCCCGATGACAGAAAGTTATTAAGTAAATATTTGAAACAATGGCCCTCAGCGATTG	1349
Qy	1586	gtagaagatacaanaaagctccgatcttatggcatalacacagaatgaaccttgaagctctg	1645
Dp	1350	GTAGAAGATACAAAAAAGTCCGATTTATGGCATACACAGATGAACCTTTAAGACTCGTG	1409
Qy	1646	aagctatcaagcatgaatcaggaatctcttggagaccttacttlaatgggaagtttgagaca	1705
Dp	1410	AAGCTATTACAGATGAATCAGGAATCTTGGAGACCTTACTTTATGGGAAATTGGAGACA	1469
Qy	1706	caatgttgattatatattlaagaatcaagcaagcagacaataacaatcacctcacgaa	1765
Dp	1470	CACCTGTGATTAATATTATTAAGATACACCAAGACAGACCATTAACATCTACCTCAGGAA	1529
Qy	1766	tcaatgagtctcgctcccttcttattcatcaggagatattcaaaaaggttgaaaaacttgaag	1822
Dp	1530	TCACGTAGTCCGCTCTTGTGATTCAAAGGAGATTACCAAAAAGGTGTAAACATTTGAAGG	1588
Qy	1826	atttccaactctgcagaggaaaatattccaatatataaatggaagcagctctgagaagatg	1885
Dp	1590	ATTTTCCAAATCTGCGCAGAGAAATATTCAAAATATATAATGAGACAGTGTAGAAAGATG	1649
Qy	1886	ggccaactaaatcagatccctcgtgtgccttgaccccgctattactagtttcgttlaatag	1945
Dp	1650	GGCCAACTAAATTCAGATCCGTGGGCTGACCCGCTATTACTAGTTTCGTTAATATG	1709
Qy	1946	agagagatctagcttcagagactatgtgcccctctctcatctgctctacaagaagatctgag	2005
Dp	1710	AGAGGATCTAGCTTCAGAGATCTATGGCCCTCTCTCCTCACTGTGCTACAAAGATCTGAG	1769

1110	ATGAAGAGCGGAAAGACTATGATATATATCTTACTGATTTCTGAAAGATGCGTCAGCT	1165
1406	ttgatalgatacaactctctctcttataccaatctcgctcagttgcgaagaacatccta	1465
1170	TTGATGATGACAACTCTCTCTCTTATTCGAAATTCGCTCAGTTGGCAAGACATCTTA	1229
1466	aaacttggtaacatacaatctgctctgaagaagagacatggagacatactgcttaagtc	1525
1230	AAACTTGGGTACATTACTTCTCTCTTAAGAGAGAGACTGGACATATGCTCCTTAGTCC	1288
1526	tcgccccgatgacagaagatataaaagtcataatttgaaacaatvgcctccagagattg	1585
1290	TGCGCCCGGATGACAGAAAGTTATTTAAAGTAAATATTGAACAATGGCCCTCAGCGGATTG	1349
1586	gtagaagatcaaaaagaatccgattatagtcatacaacaga tgaacaccttaagactctg	1645
1350	GTAGGAAGTACAAAAAAGTCCGATTATTGGCATACACAGATGAACCTTTAAGACTCGG	1405
1646	aagcatctcaagcatgaaatcagaatctcttggaccttaactttagggaaagttagagaca	1705
1410	AAGCATTTCCACATGAATCAAGAAATCTTGAGACCTTTACTTATTTGGGAAAGTTGGAGACA	1469
1706	caacttgaatatatctttaaagaatcaagaagaagagacatataaacatctaacctccagaa	1765
1470	CACCTTTGATTTATTTTAAAGATTAACCAAGCAGACCATTTATACATCTTACCTCAGGAA	1529
1766	tcaactgactgcctctcttcttattatccaagagatataccaagaagtgtlaaacaattgaag	1825
1530	TCACGATGATCGTCGTCCTTTGATTCAGAGAGATTACCAAAAGGTGTAAMAACATTGAAAG	1588
1826	atttccaattctgcgcgggagaaatatcaaatataaattggaaagtgtactgtagaagattg	1888
1590	ATTTTTCCAAATTCGCGAGGAATAATTCAAATTAATTAATGAGAGAGCTGTGAAGAAGTG	1649
1886	ggccaactcaaatcgaatccctcgatgacctggaaccgcatatacttaattctgcttaataatg	1945
1650	GSCCAACTAATTCAGATTCCTGGTGCTGACCCGCGATTACTGTAAGTTTCGTTAATATGG	1709
1946	agagagatcagcttcagagcataatggccctctctcatctgctatacaagaactctgatg	2005
1710	AGAGGATCTAGCTTCAGGAGATCTATTTGGCCCTCTCTCATCTGCTACAAAGAAATCTGATG	1769



OY	2006	atcaaaagagaacccagataatgtcgaacaagaaggaaigtcalccgtgtttctglattyg	2065
Db	1770	ATCAAAAGGGAACCAATATATGTCCAGACAAGAGGAAATGTCTCCTGTTTTCTGTATTGG	1829
OY	2066	atlgaaaaccgaagtgtgtacctcaacagagaataataaacgcttctcccacatccagtg	2125
Db	1830	ATGAGAACCGAAGCTGGTTACTTCCAGAGATATATCAACGCCTTTCTCCCATTCCAGTG	1889
OY	2126	gaatgcagctltgaagatccagagtlccaaagcctccaacatcatgcaagcatcaatgyct	2185
Db	1890	GAGTCAGCGTGTGAGATGCCAGAGTCCAAAGCCTCCCAACATCATGCAAGCATCAATGCT	1949
OY	2186	atgttttgttaagtttcgaatttcgaatttgttttcgatbaagttggcataactgfratctc	2245
Db	1950	ATGTTTTGTAGTAGTTTGTCAAGTTCTGAGTTTGTTCATGAGGTGGCATACTGGTACATTC	2009
OY	2246	taagcatltggaacaacagactgaacttcccttcgtctctctctctcgtgatatacctccaac	2305
Db	2010	TAAACATTGGAGACACAGACTGACTTCTTGTCTGTCTTCTCTGTGATATACCTTAAC	2069
OY	2306	accaaatgtgtctatlgaaagaacaactcaaccttatcccatlctcaaggagaaacgtcttca	2355
Db	2070	ACAAAATGGTCTATGAAGAACACTCACCTCACCTTCCATTCCAGATTCAGAGAAAACTGTCTCA	2129
OY	2366	tgtcgaatgaaaaaaccgaagtctcatgtgattctbggggtgccaaactcgaacttcggaga	2425
Db	2130	TGTGATGAGAAAAACCAAGGCTCTATGATCTGTGGGTGCCACACACTCAACCTTTGGANCA	2189
OY	2426	gaagcatgaacgcgcttactgaaggttcttagitgtgacaagaacactgvtatlatiaag	2485
Db	2190	GAGGCATGACGGCCTTACTGAAAGGTTTCTAAGTTTCAAGAGGGGAGGAGGACGACG	2249
OY	2486	aggcgagtlctgaagatlctca-----gcatctgtctgtgataaaca	2530
Db	2250	ACTTCTCGACCTGGAGAAAGATATTTCAGTGAAGACGACGACTACATCGACATCGTCGCA	2309
OY	2531	atgcatcttgaaccaaagaagcttctccagaattccaagaacacccctgacactgagcaaaagc	2590
Db	2310	GTCGATTGTAACCAAGAGCTTCTCCAGAAATTCAAAGACACCCCTGACACTAGGCAAAAAGC	2369
OY	2591	aattaatgcccc-----aattaatgcccc-----	2603
Db	2370	AATTATATGCCACCACAAATTCAGAAAAATGACATGAGAGACAGACCTTGTTGGCAC	2429
OY	2604	-----	2603
Db	2430	ACAGAACACTATGCCTAAATAATCAAAAATGTCTCCTCTAGTGAATTTGTGTATGCTTTGC	2489
OY	2604	-----	2603
Db	2490	GACGAGTCCCTACTCCACATGGGCTATCCTTATCTGATCTCCAAGAAAGCCAAATATGAGA	2549
OY	2604	-----	2603
Db	2550	CTTTTCTGTATGATCCATCACTCGGAGCAATAGACAGTAATBACAGCCTGTTCTGAATGA	2609
OY	2604	-----	2603
Db	2610	CACACTTCAGGCCACAGCTCCATCAACAGTGGGAGCATGTAATTAACCCCTGAGTCAGGCC	2669
OY	2604	-----	2603
Db	2670	TCCAAATTAGATTAAATGAGAAACTGGGGACAATGCGAGATCTCTTGCTTGGGATAACC	2729
OY	2604	-----	2603
Db	2730	ACTATGTACTCAGATACCAAAAAAGAAAGATGGAATTCOCAGAGAGAGTACACACAAAAAAA	2789
OY	2604	-----	2603
Db	2790	CAGCTTTAAGAAAAAGATACCATTTTGTCCCTGACGCTGTGTGAAGCAANTATGCAA	2849

[illegible]



Db	3930	AGTATTAATGGCAGCTGTACAATCTGTATCCAGGTGTTTTGAGACAGTGGAAATGTTAC	3989
OY	3656	catcaaaagcttggaaatttggcgggttggaatgaccttaatttggcgagcatctacatgcttgga	3715
Db	3990	CATCCAAAGCTGGAAATTTGGCGGGTGGAAAGCCTTATTGGCGACATCTACATCTCGGA	4049
OY	3716	tgaagcaacttttcttgggttgaagcaataagtgctagaccccccttggaaatggtctctg	3775
Db	4050	TGAGCACTACTTTTCTGGGTACAGCAATTAAGTGTACAGACTCCCTCGGAAATGGCTTCTG	4109
OY	3776	gacacatlagaatlttcagaatlacaagcttcacagagacataatgacagtgggccccaaagc	3835
Db	4110	GACACATTAGAGATTTTTCAGATTACACTTTCAGACCAATTAATGTGACGTGGCCCCCAAGC	4169
OY	3836	tggcgaacttcatatttccggatloactaatgctctggagacaccaaggagcccttctct	3895
Db	4170	TGGCCAGACTCTAATTAATTCGGATCAATCAATGCTTGACACCAAGAGCCCTTTTCTT	4229
OY	3896	ggatcaaggatggaatctctgttgacccaatggaatlacacggcatcaagaccagggtgcc	3955
Db	4230	GGATCAAGGTGGATTCGTGGTGGCACCAATGATTAATTCACGGCATCAAGACCCAGGTGCC	4289
OY	3956	gtcaagaatcttccagcctctacatctccagttatcatcatgatagtatcttgaatgga	4015
Db	4290	GTCCAGAGATTCACACCTCTACATCTCTAGTTTATCATCATGATTAATAGTCTTATGATGGGA	4349
OY	4016	agaagtggaagacttltcggagaaatctcaactcgtgaaccttaatggtctctcttggcaatg	4075
Db	4350	AGAAGTGGCAGACTTATCCGAGAAATTCACCTGACACTTAATGTCCTTCTTGGCATG	4409
OY	4076	tggaatcatctggagataaaacacaataatlttlaacctccaaatlatgtctcgatacac	4135
Db	4410	TGGATTCATCTGGGATTAACACAAATATTTTAACTCCCAATTAATTTGCTCGATACATCC	4469
OY	4136	gttttgacccaactcttatagatcttcggagcactcttgcataagatlttaatgggcgtg	4195
Db	4470	GTTTGCACCCCACTCAATTATAGATTTGCGACACTCTTCCGATGGAGTTGATGGGCTGTG	4529
OY	4196	attlaaatgtgtgcaagatgccaatgggaaatggagataaagcaataatcgaatgacaga	4255
Db	4530	ATTTTAATATGTTGCACATGCCATTTGGGANTGAGATTAAGAATAATACATATGATGCACAGA	4589
OY	4256	ttaactgtctcatcttaactcttaacaataatggtttgcacctggtctctcttcaaaagctcgac	4315
Db	4590	TTACTGCTTCATCTCTCTTTACCAATTAATGTTTGGCACTGGTCTCTTTCAAAACCTCGAC	4649
OY	4316	ttcaacctcaaaaggagagatgaatgacctgaggagacctgaagtgaataatccaanaagtgag	4375
Db	4650	TTTCACCTCCAGGAGAGAGATTAATGCCCTGGAGACTCTCAGTGATTAATATCCAAAGATGGC	4709
OY	4376	tgcaaatgtgaacttccagaagaacaaatgaaatgcaacaggagtaactactcaagggataaat	4435
Db	4710	TTCAGATGTGACTTCGAGAGACATGAAGTTCACAGAGTAACATACTACAGGAGATAAAT	4769
OY	4436	cctctgcctaacagacagatgtatgaaagaafttctcaactctcaagaaagcaagatgacatc	4495
Db	4770	CTCTGCTTACCAACATGATATGTAAGAGATTCCTCATCTCCAGCAGTCAAGATGTCATC	4829
OY	4486	agtgtgaactctcttcttcagaatcggcaaaatlaaagtttcttcaggagaaatcaaaacctct	4555
Db	4830	AGTGGACTCTCTTTTTCAGAAATGGCAAAAGTTAAAGTTTTCAGGGAATCAAGACTCCT	4889
OY	4556	ttaacactgttgttgaactctctctagaaccacaccttactgactcgttactcttcoaatcaac	4615
Db	4890	TTACACACTTGTGGAACTCTTAAGACCCACCGTTACTAGACTCGCTACTCTTCGAATTTACC	4949
OY	4616	cccagaagtttggtgtgcccacagatgaccttgaggaatgtaggtctcttgagcttgcgaagcacag	4675
Db	4950	CCCAGAGTTTGGGTGCACACAGATTTGCCCTGAGAGATGAGAGTTCTTGCGCTCGAGGCACAG	5009
OY	4676	acctctactgaactcgagcgagttc 4699	

DB	5010	ACCTCTACTGAGGGTGGCCACTGC	5033
		RESULT 9	
		US-08-558-107-1	
		/ Sequence 1, Application US/08558107	
		/ Patent No. 5910481	
		/ GENERAL INFORMATION:	
		/ APPLICANT: VOORBERG, Johannes J.	
		/ TITLE OF INVENTION: HYBRID PROTEINS WITH MODIFIED ACTIVITY	
		/ NUMBER OF SEQUENCES: 17	
		/ CORRESPONDENCE ADDRESS:	
		/ ADDRESSEE: Foley & Lardner	
		/ STREET: 3000 K Street, N.W., Suite 500	
		/ CITY: Washington	
		/ STATE: D.C.	
		/ COUNTRY: USA	
		/ ZIP: 20007-5109	
		/ COMPUTER READABLE FORM:	
		/ MEDIUM TYPE: Floppy disk	
		/ COMPUTER: IBM PC compatible	
		/ OPERATING SYSTEM: PC-DOS/MS-DOS	
		/ SOFTWARE: Patentin Release #1.0, Version #1.30	
		/ CURRENT APPLICATION DATA:	
		/ APPLICATION NUMBER: US/08/558,107	
		/ FILING DATE: 13-NOV-1995	
		/ CLASSIFICATION: 435	
		/ ATTORNEY/AGENT INFORMATION:	
		/ NAME: ISACSON, John P.	
		/ REGISTRATION NUMBER: 33,715	
		/ REFERENCE/DOCKET NUMBER: 30472/212	
		/ TELECOMMUNICATION INFORMATION:	
		/ TELEPHONE: (202)672-5300	
		/ TELEFAX: (202)672-5399	
		/ TELEX: 904136	
		/ INFORMATION FOR SEQ. ID NO: 1:	
		/ SEQUENCE CHARACTERISTICS:	
		/ LENGTH: 5035 base pairs	
		/ TYPE: nucleic acid	
		/ STRANDEDNESS: single	
		/ TOPOLOGY: linear	
		/ FEATURE:	
		/ NAME/KEY: CDS	
		/ LOCATION: 35..5017	
		US-08-558-107-1	
		Query Match 78.7%; Score 3748; DB 2; Length 5035;	
		Best Local Similarity 87.4%; Pred. No. 0;	
		Matches 4374; Conservative 0; Mismatches 60; Indels 570; Gaps 2	
QY	266	ccacacatgacaatagagctctccacctgtctcttcctgtgccttttgcatctgctta	325
DB	30	CCACCATGAGAAATAGAGCTCTCCACCCTGCTTCTTCTGCTTTTGGATTCGCTTFA	89
QY	326	gtgcaccagaagaatactactctgtgtgtgtgcagttggaactgtcatgtgacatata	385
DB	90	GTGCGACCAAGAGATACTACTGTGGTGCGAGGAACTGTCATGGAGCTATATGCAAGTG	149
QY	386	atcccggtgagctgcgcctgtgagacgcaagaatttcctctctagagtgcacaaatc	445
DB	150	ATTCGCGTGAAGCTGCGCTGTGACGCAAGATTTTCCTCTAGAGTGCACAAATCTTTCAT	209
QY	446	tcaaacacctgaagtcgtgtataaagaagacctctgtttgtatgaattcaagttt	505
DB	210	TCAACACCTTGAAGTCGTGTACAAAAGACCTGTGTTGTGAAATTCACGAGATCACCTTTTCA	269
QY	506	acatcgctgaagcaagcgacccttgatgtgtctgtagtgctcatccaatccagagct	565
DB	270	ACATCGCTAAGCCCAAGGCCACCTGTGATGGGTCTGTAAGTCTTACCATCAGGCTAGG	329
QY	566	ttatgatagaagtggtcatctaatcaacttaagaacatggtctcccatctcgttca	625



Db 330 TTATGATACAGTGTGATATACACTTAAGAACATGGCTTCCCATCTGTGAGTCTTCATG 389  
QY 626 ctgtgtgtgtactctactgtgaaagcttctgagggagctgatatgatgatacaagccagtc 685  
Db 390 CTGTGGGTATCTCTACTGGAAGCTTCTGAGGAGCTGAATATGATGATCAGACCGTC 449  
QY 686 aaagaggaagaagatgataaagtcctcccttggtggaaagccatactatgtcgtgagc 745  
Db 450 AAAGGAGAAAGAGATGATTAAGTCTTCCCTGGTGAAGCCATACATATGTCTGGGAGG 509  
QY 746 tctgtaaagagaatgtgtcccaatggcctctgacccaatgtgccttaactactacatctt 805  
Db 510 TCTGGAAGAGAAATGTCTCAATAGGCTCTGACCCACTGTGCTTACTACTCATATCTTT 569  
QY 806 ctcaatgtgagacctgtgtaaagacttgataatccagcccatgttgagccctactagtaagta 865  
Db 570 CTCTATGTGGACCTGTGTAAAGACTTGAATTCAGGCTCATTTGGCCCTACTATATAGTA 629  
QY 866 gaagaaggaagctgtgccaagaagaacacagacctgtgccaatcttataactacttttg 925  
Db 630 GAGAGAGGAGTCTGGCCAGAGAAAGACAGACACTTGCACAAATTTATACTACTTTTGG 689  
QY 926 ctgtatcttgatgaaaggaagaaagttgacctcaagaacaaagaactccttgatgcaagata 985  
Db 690 CTGTATTTGATGAAGGGGAAAAAGTTGGCAGCTCAGAAACAAGAACTCCTTGATGAGGATA 749  
QY 986 gggatgtgatctgtcgtcgggacctgacctaaatgacacagttcaatagtgtatgtaaca 1045  
Db 750 GGGATGTGATCTGTCTGGGAGCTGGGCTTAAGTCACACAGTCMAAGTATGCTAAACA 809  
QY 1046 ggtctctgcagagctgtgatgtgatgtcacaggaatcagtcatactgtgcatgtgata 1105  
Db 810 GGTCTGTCCAGAGTCTGATTTGATGATGCCACAGAAATCAGTCTATTTGSCATGTGTA 869  
QY 1106 tgggaccacctcctgaaagtgcactcaatctcctcgaaaggttcaacacatctctgtgagga 1165  
Db 870 TGGGACACCTCTCGAAGTGCACATATTTCCGAAAGGTCAACACTTTCTGTGAGGA 929  
QY 1166 accatgtgcagggggtcctgttgaaatctgcgaataaacttccctactgtctcaaacactct 1225  
Db 930 ACCATGCCCCAGGCGCTCTTGAAATCTGCCAAATTAACCTTTACGTGCTCAACACTCT 989  
QY 1226 tgaatgaccttggaacagtttctactgttctgcatatctctcccaacaacatgatgga 1285  
Db 990 TGAATGACCTTGGACACTTCTACTGTTTGTCTATCTCTTCCACCAACATGATGGA 1049  
QY 1286 tgggaagctatgtcaaaagtacagagctgtccagaggaagcccaactacgaaatgaanaata 1345  
Db 1050 TGGAACTTATGTCAAGGTAGACAGCTGTCCAGAGAAACCCCAACTCGAATGAAAAATA 1109  
QY 1346 atgaagaaggggaagactatgatgatacttaactgattctgaaatggaatgtgtgcaagt 1405  
Db 1110 ATGAAAGAGGGGAAGACTATGATGATGATCTTACTGATTTGAAATGATGATGGTCAAGT 1169  
QY 1406 ttgaatgacaaactcctctctcttataccaaatcgcactagttgcgaagaacatccta 1465  
Db 1170 TTGATGATGACAACTCTCTCTCTTTATCCAAATTCGCTAGTTGCCAABAAGCATCTTA 1229  
QY 1466 aaactgtggatcatctactgtctgtcgaagaaggaagactgtggaactatgtcccttagtc 1525  
Db 1230 AAACCTTGGGTACATTAATGCTGCTGTAAGAGAGGAGCACTGGGACTATAGTCCCTTAATGC 1289  
QY 1526 tgcgccccgagtgacagaagaagtataaagtcataatattgaaacaatgagccctcaagcgttg 1585  
Db 1290 TCGCCCCCGATGACAGATTATTAAGTCAATATTTTGAACAAATGGCCCTCAGCGGGAATG 1349  
QY 1586 gtaggaagtaacaaaaggtcgaatttatggcatcacacagatgaaccccttaagactgtg 1645  
Db 1350 GTAGGAGATGCAAAAAAGTCCGATTTATGGCATACACAGATGAACCTTTAAGACTGTGG 1409  
QY 1646 aagcatctcagcatgaacgaagactcttggaaccttaacttlatgggaagtgtgagaga 1705  
Db 1410 AAGCTATTCAGCATGAAATTCAGGAATCTTGGACCTTATTTATGGGAGAGTTGAGAGACA 1469

QY 1706 cactgtgatataatgaagaatcaagcaagaccataaacatctaccctcaagga 1765  
Db 1470 CACTGTGATTTATATTAAGAAATCAACGAAGACCATATTAATCTACCTCAGCGAA 1529  
QY 1766 tcaactgtatgcgtctctgtatlatcaagagatlaaccaaaagttgtaaacatttgaaag 1825  
Db 1530 TCACTGATGTCCGTCCTTTGATTCAGAGAGATTTACCAAAAAGGTAAACATTTGAAG 1589  
QY 1826 atttccaatctgtgcagaggaataltcaataataaavgacagtgtctagaagatg 1885  
Db 1590 ATTTTCCAAATTTGCCCCAGGAATAATTAATAATGAGACAGTACTGTGAAGATG 1649  
QY 1886 ggcacaataatgaatccctcgtgtgcctgagcccgctatctctactagtttgtaataatg 1945  
Db 1650 GGCCAACTAATCAAGATCTGTGGTGTGACCCGCTATTAATCTGATTTGTTAAATATGG 1709  
QY 1946 agagagatccagcttcaagactcatgtgcctctccctcatctgtcaaaaagatctgtag 2005  
Db 1710 AGAGAGATCTAGCTTCAGAGACTATTGGCCCTCTCCATCTGCTACAAAGAACTGTAG 1769  
QY 2006 atcaaaaggaagaaccagataatgtcagaacaaaggaatgtcatcctgttctgtatctg 2065  
Db 1770 ATCAAAAGGAAACCAAGATATGTCAAGACAAGAGAAATGTCAATCTGTTTCTGTATTG 1829  
QY 2066 atgaagaagcaagctgtgtactcctcaagagaaatatacaagccttctcccaatccagctg 2125  
Db 1830 ATGAGAACCCAGAGCTGTACTCTCAAGAAATATACAACTTTCTCCAAATCCAGCTG 1889  
QY 2126 gaatgcagcttgaggaatccagagcttcccaagcctcccaacatcatgacagatcaatgct 2185  
Db 1890 GAGTGCAGCTTGAAGATCCAGAGTCCAAAGCTCCAAATCATGACAGCATCAATGCT 1949  
QY 2186 atgtcttctgtagtctgtcagttgtcagttgtgttgatgaggtgtgcactgtgtacatctc 2245  
Db 1950 ATGTTTGTGATGATTTGGCACTTGTCAATGTTGTTGATGAGTGGAGCTATCTGTATTC 2009  
QY 2246 taagcatggaagcaagagagacttctctgtctctctctctctgtatataactcaaac 2305  
Db 2010 TAAAGCATGGAGCACAGACTGACTCTCTTCTGTCTCTCTCTGTGATTAATCTTCAAC 2069  
QY 2306 acaaaatgtgtctatgaagaacacatcccatctccatctcaagagaagactgtgtctca 2365  
Db 2070 ACAAAATGCTTATGAAGACACACTGACCTATTCCATCTCAGAGAAACGTCTTCA 2129  
QY 2366 tctogatggaaaaaccagagctctatgtgattctgggtgtgcacaaactcaagacttgcgaga 2425  
Db 2130 TGTGATGGAACCAACCCAGAGCTATGATGATCTGGGGTCCCAACTCAGACTTTCGGACA 2189  
QY 2426 gaagcatgaccgctactgaaggttcttagttgtgacaagaacacactggtgatattacg 2485  
Db 2190 GAGGCATGACCGCTTACTGAAGGTTTCTAAGTTGTATTCAGAGGGGGAGAGACACG 2249  
QY 2486 aggaagttatgaagatattca-----gcatactgtgtagtaaaaaca 2530  
Db 2250 ACTATCTGACCTGAGAAATATTTCACTGAAGACAGACGCTACATGCAATGTGTGACA 2309  
QY 2531 atgccaatgaaacagaagctctctccagaatccaagaacccctagcaataggcaaaagc 2590  
Db 2310 GTCTGATTTGAACCAAGAAAGCTTCTCCAGAAATCAAGACACCTAGACATGAGCAAAAGC 2369  
QY 2591 aatttaatgcac----- 2603  
Db 2370 AATTTAATGCAACACCAATTCAGAAAATGACATGAGAAAGACTGACCTTGGTTGAC 2429  
QY 2604 ----- 2603  
Db 2430 ACAGAACACTATGSCCTAAATATCAAAATGTCTCCTGTATGATTTGTGTATCTCTTGC 2489  
QY 2604 ----- 2603  
Db 2490 GACAGATCTCTACTCACATGGGCTATCCTTATCTGATCTCAAGAAGACCAATATGAGA 2549



QY 2604 ----- 2603  
Db 2550 CTTTTCGATGATCCATCACCCTGGAGCAATAGACAGTAATTAACAGCCTGTCTGAATGA 2609  
QY 2604 ----- 2603  
Db 2610 CACACTTCAGGCCACAGCTCCATCATCAGCTGGGACATGGTATTTCACCCCTGAGTACGGCC 2669  
QY 2604 ----- 2603  
Db 2670 TCCAAATTAAAGTTAAATGAGAAACTGGGACAACTGCAGATCCTCTGTGGATTAACC 2729  
QY 2604 ----- 2603  
Db 2730 ACTATGACTACTAGATACCAAGAAGAGTGAATCCCAAGAGAATGCCAGAAAAA 2789  
QY 2604 ----- 2603  
Db 2790 CAGCTTTAAGAAAAAGATACCATTTTGTCCCTGACGCTTTGTGAAGCAATCATGCCA 2849  
QY 2604 ----- 2603  
Db 2850 TAGCAGCAATAATAGGGGACAAATTAAGCCGGAATGAAGTCACTGGGCAAAAGCAAG 2909  
QY 2604 ----- 2635  
Db 2910 GTAGAGCTGAAGGCTGTCTCTCAAAACCCACAGCTTGTGAAGCCATCAACGGGAAA 2969  
QY 2636 taactgtaactctctcagatcaagaaggaaattgacatgatataccatataag 2695  
Db 2970 TAACTCGTACTCTCTCAGTCAGATCAAGAGGAATTGACTATGATATACCATATACG 3029  
QY 2696 ttgaatgaagaagaagatttgcacattatgatgatgaatgaagaatcagagcccgca 2755  
Db 3030 TTGAAATGAAGAAGAGATTGTCATTTATGATGAGATGAATGAATCAGAGCCCGCA 3089  
QY 2756 gcttccaaagaacaacgacataatttattgctgcagtggaagagctctggaattatg 2815  
Db 3090 GCTTTCAAAACAAAACAGCACATATTTATTTGCTGCAGTGAAGGCTGTGGGATTTATG 3149  
QY 2816 gaatgagtagtctcccaatglttccaaagaacaggtcgaagtgagcagtgccctcagt 2875  
Db 3150 GGATGAGTAGTCCCCCAGCATTTTCTAAGAAACAGGGCTCAGAGTGGACATGCTCCAGT 3209  
QY 2876 tcaagaagtgttcttccagaatcttactga tggctccttactcagccctataacgltg 2935  
Db 3210 TCAAGAAGTTGTTTCCAGGAATTTACTGATGGCTCCTTACTCAGCCCTATACCGTG 3269  
QY 2936 gagaactcaaatgaacaatttggagctcctgggccaataataagacagaagtgtgaata 2995  
Db 3270 GAGAACTAAATGAACATTTGGGACTCCGTGGGCCATATATTAAGACACAGAATTTGAAGATA 3329  
QY 2996 atatcagtgaacttccaagaatcagagctctgctccctattccttcatcttaacctta 3055  
Db 3330 ATATCATGTAACCTTTCAAGAAATCAGGCCCTCTGCTCCATTTCTTCAATTTACCTTA 3389  
QY 3056 ttctctta tgaaggaagatcagaagcagaagcagaacctagaaaaaacttltg tcaagccta 3115  
Db 3390 TTTCCTTGAAGGAGATCAGAGGCAAGACGAACCTAGAAAAAACTTTGTCAAGCCCTA 3449  
QY 3116 atgaacccaataacttacttcttggaaagtgaacaatcataatgaccaccaataagaatgag 3175  
Db 3450 ATGAACCAAAACTTACTTTTGGAAAGTGAACATCATATGSCACCCCACTAAAGATGAGT 3509  
QY 3176 ttgactcaagaagcctgggctattcttctga ttttgaaccttggaaaaaagtgtgacctag 3235  
Db 3510 TTGACTCAAAAGCCTGGGCTTATTTCGTGATGTGACCTGAAAAAATGTGCACTAG 3569  
QY 3236 gctcgtatggacccctctctgttctgcacaactaaacactgaacctgtcatalggaagac 3295  
Db 3570 GCGTGATTTGAGCCCTTGTGTGTGCACACTTAACACACTGAACCTCTCTCATGGGAAC 3629  
QY 3296 aagtgaagtaacaggaattgtctctgttcttcaacatcttga tgaagacaaaagctgtg 3355

Db 3630 AAGTACAGTACAGGAATTTGCTGTGTTTTCACCATTTTGAAGACCAAAAGCTGGT 3689  
QY 3356 acttcaactgaaata tggaaagaactcgaaggtcccttgcataatccaaga tggaaatc 3415  
Db 3690 ACTTCACTGAAATATGGAAGAAACCTCAGGGCTCCCTGCAATATCAGATGGAATATC 3749  
QY 3416 ccactttaaagaagaattatcgcttccatgtcaatcaatcagctacataatgatactactac 3475  
Db 3750 CCACCTTTAAAGAGATTAATGCGTTCATGCAATCAATAGGCTACATATATGATGACTAC 3809  
QY 3476 ctggcttagtaattggtcgaagatcaagaagtatgatatcgtcctaagcatgggca 3535  
Db 3810 CTGGCTTAGTAATGGCTCAGATCAAGGATTCAGATGATCTCTCAGCATGGGCGCA 3669  
QY 3536 atgaacacatccatcttacttcaattcag tggacatgtgtcaactgtacgaaaaaagag 3595  
Db 3870 ATGAATAACATCATTTCTTTCATTTCACTGACATGTGTTACTGTACGAAAAAAGAG 3629  
QY 3596 agataaaatgacatgtacaaatctctatccaagtglttlttgagacag tggaaatgttac 3655  
Db 3930 AGTATTAATATGGCACTGTACATCTATCCAGGTGTTTGAAGACAGTGAATGTATAC 3689  
QY 3656 catccaagctggaaatttggcgggtggaa tgccttatttgcagacatcctacatgtggaa 3715  
Db 3990 CATCCAAAGCTGGAAATTTGGCGGGTGAATGCTTATTTGGGACATCTACATGCTGGGA 4049  
QY 3716 tgaagacacttttctgtgtgtaagcaaatgaatttcaagaccccttgggaatgctctgt 3775  
Db 4050 TGAGCACCTTTTGTGTGTGATGACAAATAGTGCACACTCCCTGTGAATGTGCTGTG 4109  
QY 3776 gacacattagagatttlaagatlacagcttcaagacaaatgaacag tgggccccaaagc 3835  
Db 4110 GACACATTAGAGATTTTCAGATTTACAGCTTCAGGACATATGACAGTGGGCCCAAGC 4169  
QY 3836 tggcgaacttcaattatctcgagatcaatcaatgccttggagcacaagagagccttctc 3895  
Db 4170 TGGCAGACTTCATTATTTCCGGATCAAAATCAATGCTGTGAGCACCAAGACCCCTTTCTT 4229  
QY 3896 ggaatcaagtgatctgttggcaccaatgtatcttcaagcatcaagaaccagaagtgc 3955  
Db 4230 GCATCAAGGTGATCTGTGGCACCAATGATTAATTCAGGACATCAAAACCCAGGTC 4289  
QY 3956 gtcaagaattctccaagcctcatalctctcagttatcatalcatatgatatgttgaatga 4015  
Db 4290 GTCAAGATTTCCAGCCTCTACATCTCTCAGTTATATCATATGATATGATGATGGA 4349  
QY 4016 agaagtggcaactatctcgaaggaattccacttggaaaccttaagtgtcttcttggcaatg 4075  
Db 4350 AGAAGTGGCAACATTATGAGGAATTCACCTGGAACCTTAATGCTCTTTGGCAATG 4409  
QY 4076 tggattctctgggataaacaacaatatltttaaacctccaatatgtgcgataatacc 4135  
Db 4410 TGGATTTCTCTGGGATTAACACATATTTTAAACCTCCATATTTATGCTGATACATCC 4469  
QY 4136 gtttgcaccaactcaatataatgcatctgcagcactcttgcagatgag tlgatgggctgtg 4195  
Db 4470 GTTTGCACCAACATATATATGATTCAGACACTCTTGGATGAGTGTATGGGCTGTG 4529  
QY 4196 atttaaatagttgcaagcatgccaattgggaatggagatgaagaataatcagaatgacaga 4255  
Db 4530 ATTTAAATAGTTGAGCATGTCATTTGGAAATGAGAGATTAAGCAATATCGATGCAACGA 4589  
QY 4256 ttactgtctcaacttcaacttcaaatatgttgcacactgtgtcccttcaaaagctgac 4315  
Db 4590 TTACTGCTTCACTCTACTTTACCAATATGTTTGGCACCTGCTCTCTTCAAAAGCTGCAC 4649  
QY 4316 ttcaactccaagggagagtaatgtccttggagaccctcaggtgaaataatccaaaag tggc 4375  
Db 4650 TTCAACCTCAAGGGAGGATATGCTGTGAGACCTCAGGATTAATCCAAAAGAGTGGC 4709  
QY 4376 ttcaagtgtacttccagaagaacatgaagaatcacaagatgaactactcagggagtaaat 4435



Dh 4710 TCGAAGTGAAGTTCAGAGCAATGAAAGTACAGGAGTAACACTACTCAGGAGTAATAAT 4769  
Qy 4436 ctgcgtaccagatgtatgtgaagaagtcttcctcatctcagcagcagatggccatc 4495  
Dh 4770 CTCGCTTACAGATGATGTGAAGAGTTCCTCATCTCCAGCAGCAAGATGGCCATC 4829  
Qy 4496 agtgaactctcttcttcagaatgcaaaatgaagttcttcaggaggaatcaagactcc 4555  
Dh 4830 AGTGAAGTCTCTTTTTCAGATGCGAAAGTAAGGTTTTCAGGGAAATCAAGACTCCT 4889  
Qy 4556 tcacacctgtgtgaactctctcagaccaccgttactgactcgtactcgaatcacc 4615  
Dh 4890 TCACACCTGTGGTGAAGTCTCTTAACCCACCCGTTACTGACTCGCTTACCTTCGAATTAC 4949  
Qy 4616 cccagagttgtgaccagaattgccctgaagatgaggttctggctgaggaagcagc 4675  
Dh 4950 CCCGAGTGGGAGTGCACAGATGGCCCTGAGAGTGAAGGTTGCGGCTGCGAGGCACAGG 5009  
Qy 4676 acctactactcagcagagctc 4699  
Dh 5010 ACCTCTACTGAGGTGGCCACTGC 5033

## RESULT 10

US-09-243-539-1  
; Sequence 1, Application US/09243539  
; Patent No. 6130203  
; GENERAL INFORMATION:  
; APPLICANT: VOORBERG, Johannes J.  
; TITLE OF INVENTION: HYBRID PROTEINS WITH MODIFIED ACTIVITY  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/243,539  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/558,107  
; FILING DATE: 13-NOV-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: ISACSON, John P.  
; REGISTRATION NUMBER: 33,715  
; REFERENCE/DOCKET NUMBER: 30472/212  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)672-5300  
; TELEFAX: (202)672-5399  
; TELETYPE: 904136  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 5035 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 35..5017  
; US-09-243-539-1

Query Match 78.7%; Score 3748; DB 3; Length 5035;  
Best Local Similarity 87.4%; Pred. No. 0;  
Matches 4374; Conservative 0; Mismatches 60; Indels 570; Gaps 2;

Qy 266 ccacacatgcaatatagagctctccaccctgctcttccttgcttgcccttttggaattctgttta 325  
Dh 30 CCACCATGGAATATGAGCTCTCCACCTGCTTCTTCTGCTCTTTGGCGATTTGCTTTA 89  
Qy 326 gtgcacacgaagaatactactcgtgtgacgtggaactgcatatggaactatgcaagt 385  
Dh 90 GTGCACCAAGATPACTACTGGGTGCATGGAACGTGTATGGAGCTATATGCCAAGTG 149  
Qy 386 atctcgttgagctgctctgtagcgaagattctcctcctagagtgcgaatcttccat 445  
Dh 150 ATCTGGGTGAGCTGCTGTTGAGCGCAAGATTCTCTCTAAGTGCAGCAAAATCTTTTCAT 209  
Qy 446 tcaaacacctgaactcgtgttaaaaaaagactgtttgtgaagattcaagcttcttca 505  
Dh 210 TCACACCTCTAGTGTGTACAAAAGACTGTGTTGTAGATTCCAGGATCACCTTTTCA 269  
Qy 506 acatcgtctaagccaagccaccctgtagtggctgtcctcctacatccaggctgag 565  
Dh 270 ACATCGGTAGCCCAAGGCCACCTGGATGGTGTGCTGATGCTTACCATCCAGGCTGAGG 329  
Qy 566 ttatgatatacagtgtcatttaacattaaagaacatggcttccatccctgtaagctcatg 625  
Dh 330 TTTATGATACAGTGTGATACATTAAAGAACATGGCTTCCATCCCTGTCAGTCTCATG 389  
Qy 626 ctgttggtgatactctactggaagaattcttgagagagctgaatatgatacagaacagtc 685  
Dh 390 CTGTGGGTATCTCTACTGGAAGCTTGTGAGGAGCTGAATGTGATGATCAACACAGTGC 449  
Qy 686 aaggaggaagaagaattgaagaattccctgttggaacccaatacattgtctggcag 745  
Dh 450 AAGGGAAGAAAGATGATGAATGAATGCTCCCTGTGTGAACCATATATATGTGTGGCAGG 509  
Qy 746 tccctgaagaagaatggtccaaatggcctctgacccactgtgcttactactatctt 805  
Dh 510 TCCCTGAAGAAGATGGTCCAAATGGCTGTGACCCACATGCTTACTACTATATCTTT 569  
Qy 806 ctatctgagaccctggttaaaagacttgaatcaaggcttcattggaagccctactgtatgta 865  
Dh 570 CTGATGTGACCTGTGTAAAGACTTGAATTCAGGCCCTTGAAGCCCTACTATGATGTA 629  
Qy 866 gaagaaggagctctgccaagaagaacacagacctgacaaattatactactttg 925  
Dh 630 GAGAAGGAGACTTGGCCAAAGAAAGACAGACCTTGACCAAAATTATATCTCTTTTG 869  
Qy 926 ctgtatttgtagaagggaagaattggaactcagaacaagaactcctgtatgacagata 985  
Dh 690 CTGTATTTGATGAAGGGAAGAAAGTTGGCAGTCAGAAACAAAGAACTCTTGATGCAGATA 749  
Qy 986 ggaatgctgcatctgctcgggacctgacctaaatgcaacagacgaatggtatgttaaca 1045  
Dh 750 GGGATGCTGCATCTGCTCGGGCCCTGGCTTAAATGACACAGTCAATGGTTATGTAAACA 809  
Qy 1046 ggtcctgcgaagctcgtatgtagtgcacaggaatcagtcctatgtgcatgtatggaa 1105  
Dh 810 GGTCTGTGCCAGGTCTGATTTGATGGCACAGAAATCAGCTATTTGGCATGTATGGAA 869  
Qy 1106 tgggacacactcctgaaatgacactcaatatctcctgaaaggtcacaacttctgtgagga 1165  
Dh 870 TGGGACACACTCTGAAGTCACTCAATATTCCTCAAGAGTCAACATTTCTGTGAGGA 929  
Qy 1166 accatgcgaaggtgctcttggaatctgcgaataacttcttactgtcctcaaacactc 1225  
Dh 930 ACCATGCGCAGGCGCTCTTGGAAATCTCGCAATTAATCTTCTTACTGCTCAACACTCT 969  
Qy 1226 tgaatgaccttgagacagtttctactgtttgtcatactcttccacacacatgtaggca 1285  
Dh 990 TGATGACCTTGGACAGTTTCTTACTGTTGTGATATCTCTTCCACCAAGATGATGGCA 1049  
Qy 1286 tggaaagcttatgtcaaaagtagacagctgtccagaaggaaacccaactagaatgaaata 1345  
Dh 1050 TGGAAAGCTTATGTCAAAAGTAGACAGCTGTCCAGAGGAACCCCAACTATGCAATGAAAAATA 1109



QY	1346	agaagaagcggaagcetaigatigatcttactgattctgaaatgagtgtgltcagtc	14050
Db	1110	ATGAAGAAGCGGAAGACATATGATGATATCTTACTGATTTCTGAAATGATGTGGTCAGCT	1169
QY	1406	tgtatgatgaacactccctccctcttccaaacttgctcagtttgcgaagaagatcccta	1465
Db	1170	TTGATGATGCAACTCTCTCTTCTTTATCCAAATTCGCTCAGTTGGCCAGAGAGATCTCTA	1229
QY	1466	aaacttggtagatactaatgtctgtctgaagaagagagacttggagacttgcctctagtc	1535
Db	1230	AAACTGGGTAGCAATTACATTGCTGCTGAGAGAGAGACTGGGACTATGCTCCCTTAGTGC	1289
QY	1526	tcgcccccgatgacagaagtataaagaatcaatacttgacaataatggccctaacgcgattg	1555
Db	1290	TCGCCCCCGATGACGAGAGATTATAAAGTCAATATTTTGAACAATGGCCCTCAGGGAGTTG	1349
QY	1586	gtgaggaagtacaaaagtcgcatttatgccaacagatgaacactttaagctctgtg	1645
Db	1350	GTAAGGAATGACAAAAAGTCCGATTTATGGCATACAGATGAATACCTTTAAGACTCGTG	14090
QY	1646	aagcatttcagcatalgaatlcaggaatccttggagccttacttlatlggggaagtltgagaca	1705
Db	1410	AAAGCTATTGACATGATCAGGAATCTTGGACCTTACTTATGGGGAAGTTGAGACA	1469
QY	1706	caactgttatataattaaagaatcaagcaagacatatatactaccctcaagaa	1755
Db	1470	CAGTGTGATTATAATTTAAATAACCAAGCAACACATATAACATTTACCTTCAGGAA	1539
QY	1766	tcactatgtccgtccctctgtatccaagagattaccaaaagtgctaaacatttgaag	1825
Db	1530	TCAGTATGTCCGTCCTTGTATTCAAGGAGATTACCAAAAAGTGTAACATTTTGAAGG	1589
QY	1826	atttccaaltctgcagagaaatalccaataataatgagacatgacttgaagaatg	1885
Db	1590	ATTTTCCAAATTCGCCAGAGAAATATTCAAAATATAATGACAGTAGCTAGAACATG	1649
QY	1886	ggccaactaaatagatcctcggtgctgacccgctattacttagtttgttaatatgt	1945
Db	1650	GGGCAACTAAATAGATCCCTCGGCGCTGACCCGCTATTACTGATTTCTGTAAATATGG	17090
QY	1946	agagagatctagcttcaagacataatggccctcccaactgtgttaaagaactcttag	2005
Db	1710	AGGAGATCTAGCTTCAAGGACTCATTTGGCCCTCTCTCATCTGTGTACAAAGAACTGTAG	1769
QY	2006	atcaaaagaggaacacagataatgtcagacaagagaaatgtcalactctgttttcgttattgt	2065
Db	1770	ATCAAAAGAGAAACAGATATGTGCAGACAAGAGGAATGCATCTGTGTTCTGTATTTTG	1829
QY	2066	atgagaaacgaagctgtgtacctcaacagaaatatacaagccttctcccaatccagctg	2125
Db	1830	ATGAGAACCGAAMCTGTGTACCTCACAGAGATATATACAGCCTTTCTCCCATCTCACCTG	1889
QY	2126	gagtgagacttggagatccagagtttccaagcctcccaacatcatatgcagacataatgct	2185
Db	1890	GAGTGAAGCTTGAGGATCCAGAGTTCCAAAGCCTTCCAAATCATATGCACAGATCATAGGCT	1949
QY	2186	atgtttttagatagtlttgcagttgtlcaagtttgttctgtctgaatgagtgagataactc	2245
Db	1950	AGTGTTTTGAATGATTGCAATTTGCAATGTGTTGTTGATGATGAGGTGCADATGTAATTC	2009
QY	2246	taagcatgtgagacacagactgtaactctctctgtctctctctctctgtataacttcaaac	2305
Db	2010	TAAAGCAATTGGAGACAGACAGACTTCCTTTCTGTCTTTCTCTCGAGTATAACCTTCAAC	2069
QY	2306	acaaaaatgtctatgaagacacatacccatcttcccatcttcaaggagaacttcttca	2365
Db	2070	ACAAAAATGCTATGAAAGACACACTCACCTATTTCATTCACGGAATCTGTCTTCA	2129
QY	2366	tgtctgatgaaacccagctctatgtatcttgagggtgcacaaactcagacttgcagaca	2425
Db	2130	TGTGCATGTGAAAACCCAGGCTATGATTTGTGGGGTGCACAACTCAGACTTTGGGAAACA	2189
QY	2426	gaagcatgacccgctacttgaagtltctagtltgtgaacaagaacatgtglatbatactg	2485

Db	2190	GAAGCATGACCGCCTTACTGTGAAGGTTTCTAGTTGATTTCCAGAGGGGGGAGGACGACG	2249
Qy	2486	aagacagtlatgaaatalttcca-----gcatactgcctgagtaaaaca	2530
Db	2250	ACTATCTGGACCTGGAGAGATATTTCAGTAAAGCAGCAGCTCATCGACATCTGTGACAA	2309
Qy	2531	algcactgaaaccaagaagcttctccagaatlcgaagcaacctagcactaggcaaaagc	2590
Db	2310	GCTGATTGAACCAAGAAGCTTCTCCGAAATTCAAGACACCCCTAGCCTAGGCAAAAAGC	2369
Qy	2591	aatttaatgcccac-----	2603
Db	2370	AATTTAATGCCACACCAATTCCAGAAAATGACATAGAGAAGACTGACCCTTGTTGCAC	2429
Qy	2604	-----	2603
Db	2430	ACAGAACACCTATGCTTAATAATACAAAATGTCCTCTAGTATTTGTTGATGCTCTTGC	2488
Qy	2604	-----	2603
Db	2490	GACAGAGTCTCACTCCACATGGGCTATCTTATCTGATCTCCAGAAGCCAAATATGAGA	2549
Qy	2604	-----	2603
Db	2550	CTTTTCTGATGATCCATCACCTGGAGCAATGACAGTAATAACGCCGTCTGAATGA	2609
Qy	2604	-----	2603
Db	2610	CACACTTCAAGGCCACAGCTCCATCACAAGTGGGAGCATGATTTACCCTGATCAAGCC	2669
Qy	2604	-----	2603
Db	2670	TCCAAATTAGATTAAATGAGAAACCTGGGAGCAACTGCAGATCCTCTTGCTTGGATTAAC	2729
Qy	2604	-----	2603
Db	2730	ACTATGTACTCAGATACCAAAAGAGAGTGAATCCCAAGAGAGTCAACCAAAAAA	2789
Qy	2604	-----	2603
Db	2790	CAGCTTTTAAGAAAAAGATACGATTTTGTTCCTGAAGCGTTGTGAACATTCATGCA	2849
Qy	2604	-----	2603
Db	2850	TAGCAGCAATTAATGAGGAGCAAAATAGCCCAATAGAGTCACTGGGCAAAAG	2909
Qy	2604	-----	2603
Db	2910	GTAGACTGAAGAGCGCTGTGCTCTCAAAACCCACACTTTGAAACCCCTCAACGGGAA	2969
Qy	2636	taactcgtactactcctcgaatcgatcgaagcaagaaatgactatgtaacacatcag	2695
Db	2970	TAACTCGTACTACTCTTCAGTCAGATCAAGAGAAATGACATATGATACCATATCTAG	3029
Qy	2636	ltgaaatgaagaagaagatlttgacattatgatagtgagatgaaatcagagcccccga	2755
Db	3030	TTGAATTAAGAAAGAAAGATTTTACATTTATGATGAGATCAAAATTCAGAGCCCCGCA	3089
Qy	2756	gcttccaaaagaacccgcacactatttatctgtcgaatggaagagcctctggagttatg	2815
Db	3090	GCTTTCAAAAACAAACACGACACTATTTTATCTCTCAGAGGAGAGCTCTGGGATATAG	3149
Qy	2816	gagatgagtagctcccaacabttcttaagaacaggctcgaagtgcagtgcctccagt	2875
Db	3150	GGATGATAGCTCCCCACATGTTCTTAAGAAACAGGGCTCAGAGTGGCACTGCTCTCAGT	3209
Qy	2876	tcagaagaagtgcgttccacagaatttactgatagtctccttactcagccctataccg	2935
Db	3210	TCAAGAAAGTTGTTTCCAGGAATTAATCTATGGTCCCTTACTCAGCCCTTAATACGTG	3269
Qy	2936	gagaactaaatgaaacatttggagctccctgggcccataataagacagaagtltgaata	2995



Dh 3270 GAGAACTAAATGAACTTTGGAGCTCCTGGGGCCATATATAGAGCAGAAGTTGAAGATA 3329  
Qy 2296 atarcatgtaacttcaagaatcaggcctctgctcccttacttcttactttagcctta 3055  
Dh 3330 ATATCATGTAAGTCTTGAAGATAGGCTCTGCTCCCTATCTTCTTATCTTAGCCCTTA 3389  
Qy 3056 ttctcttctgaagaatcagaagcagaagcagaacctaagaacattgtcgaagccta 3115  
Dh 3390 TTTCTTATGAGGAAGATCAGAGCAGACAGACCTAGAAAAAATTTGTCAAGCTTA 3449  
Qy 3116 atgaacacaaacttacttcttgaagaatgcaacatacatatgcaacccactaagaatgagt 3175  
Dh 3450 ATGAAACCAAACTTACTTTTGGAAATGCAACATCATATGACACCCACTAAAGATAGT 3509  
Qy 3176 tttagctgcaagcctggtctatttctctgagtgttgaccctggaaaaaagatgctaccag 3235  
Dh 3510 TTGACTGTCAAGCTGGGCTTATTTCTGTATGTGACCTGGAAAAAGATGTGCACACAG 3569  
Qy 3236 gctgtaattgacccctctgctctgcaactaacaacactgaaacccctgctcatgaggagac 3295  
Dh 3570 GCCTGATGGAACCCCTTCTGTGCTGCTGCACACTAACACACTGMAACCTGCTCATGGAGAC 3629  
Qy 3296 aagtgacagtaacaggaattgtctctgttttccacacatcttgatgagacccaaagctggt 3355  
Dh 3630 AAGTGACAGTACAGGAATTTGCTGCTTTTTCACATCTTTGATGAGACCAAAAGCTGCT 3689  
Qy 3356 acttctctgaaataatggaagaagaactgcagagctccctgcataatccagaatggaagatc 3415  
Dh 3690 ACTTCATGAAATATGGAAGAAAGAAATGCAAGGCTCCCTCAATATCCAAATGGAATATC 3749  
Qy 3416 ccaactttaaagaagaattatcgcttccatcaatcaatgtaactaataatgataactatc 3475  
Dh 3750 CCACCTTTTAAAGAGANTATACGCTTCATGCAATCAATGCTCATATATGATACACTAC 3809  
Qy 3476 ctggtctagtaatgctcgaagatcaagaatctcgatggtatctgtctcgaatgagca 3535  
Dh 3810 CTGGCTTAGTAATGCTCAGATCAAAAGATGATGATGCTATGCTCAGCATGGGAGCA 3869  
Qy 3536 atgaacacatcattatcattcattgaatgagatgcttctactgttagagaaaaaagag 3595  
Dh 3870 ATGAAACATCATCTTATTCATTCATGATGAGATGCTGCTGCTGCTGCTGCTGCTGCTG 3929  
Qy 3596 agataaaatgagcagtaacatctcatcagctgcttcttgagacagtgtaaatgttac 3655  
Dh 3930 AGATATAAATGAGCAGCTGTAACATCTCATCCAGCTGTTTGAAGACAGTGAATGTTAC 3689  
Qy 3656 catcacaagctggaatttggcgggtggaatgctcttatttggcgaacatcatatgctgga 3715  
Dh 3990 CATCAAAAGCTGGAATTTGGCGGTGGAATGCTTATTTGGCAGCATCTCATGCTGGGA 4049  
Qy 3716 ttgagcagacttctctgtgtgtacagcaataatgtcagaactcccttgggaatggtctctg 3775  
Dh 4050 TGAAGCACACTTTTCTGCTGTACAGCAATGATGCTGAGACTCCCTGGGAATGGCTTCTG 4109  
Qy 3776 gacacattagaatttcaagattacagctcagaacaataatgaagctgggcccccaag 3835  
Dh 4110 GAACATTAAGATTTTCAATTAACAGCTTCAGACATATGACAGTGGGCCCAAGC 4169  
Qy 3836 ttggcgaactcatatctccgatacaatcaatgactgagcaccagaagaccccttctct 3895  
Dh 4170 TGGCCAGACTTCTATTTCCGGATCAATCAATGCTGAGACCAAGAGACCCCTTTTCTT 4229  
Qy 3896 ggaataagtgatctgtgtgagcacaatgataatcaagcacaagaagcgaagtgccc 3955  
Dh 4230 GGATCAAGGGGATGCTGTGGACCAATGATTTATCAAGGATCAAGACCAAGGATGCC 4289  
Qy 3956 gtcaagaattctcagcctcatatcctcagttatcatcatgataagctgagtgga 4015  
Dh 4290 GTCAAGAGTTCTCAGAGCTTACATCTCTAGTTATCATCATGATAGCTTGATGGGA 4049  
Qy 4016 agaagtgagcagactatcagaagaatccacatggaacctaagtcttcttggcaatg 4075  
Dh 4350 AGAAGTGGCAGACTTATGAGGAATTCACCTGGAACCTTAATGCTTCTTGGCAATG 4409

Qy 4076 ttgattcatctgggataaacaacataattttaaccctcccaattattgctcgatactcc 4135  
Dh 4410 TGGATTCTCTGGATATAAACACATATTTTAACTCCCTCAATATTGCTGATACATCC 4469  
Qy 4136 gtttgcacccaactcatatagatcatctgcagcactcttcgacagtgatgtaatgagctgtg 4195  
Dh 4470 GTTTGCAACCCACATTAATAGCATTTGCGAGCACTCTTCGATAGGATTTATGAGGCTG 4529  
Qy 4196 atttaaatgttgcagatgcaatggaatggaatggaatggaatggaatggaatggaatgga 4255  
Dh 4530 ATTTAAATAGTTGAGCATGCTCATTTGGGAATGAGATGAAGATTAAGCAATATCATGAC 4589  
Qy 4256 ttactgctcatctacttcaacataatgtttgacacactgtctcccttcaaaagctgac 4315  
Dh 4590 TTACTGCTTATCTCTATTTTCAATATGTTTGGCACCTGCTGCTTCAAAAGCTGAC 4649  
Qy 4316 ttcaactccaaggagagatgtaatgcttgagacactcagtgtaataatccaaagatgagc 4375  
Dh 4650 TTCACTCCCAAGGAGAGATATGCTGAGACTCTGAGTGAATATCAAAAGATGGC 4709  
Qy 4376 ttgaagtgcacttccagaagaatgaaatgcaagaagtaactactcaagtgagtaaat 4435  
Dh 4710 TGCAAGTGACTTCCAGAACATGAATGAAGTCAAGAGTAACTACAGGAGATAAAT 4769  
Qy 4436 ctctgcttaccagatgatagtgaagagttctctcatctccagagatgcaatgagcctc 4495  
Dh 4770 CTCTGCTTACAGATGATATGTAAGAGATTCCTCATCTCAGACATCAAGATGGCCATC 4829  
Qy 4496 agtgcagctccttcttcaagaatgcaagaatgaaatgaaatgaaatgaaatgaaatgaaatg 4555  
Dh 4830 AGTGAATCTCTTTTTCAGAAATGCAAAATGATTAAGTTTTCAGGAAATCAAGACTCT 4889  
Qy 4556 tcaacactgtgtgtaactctctagaccaccogtactgactgcttactcgaatctacc 4615  
Dh 4890 TCACACCTGTGTAAGTCTCTAGACCCAGCTGACTGCTGCTGCTGCTGCTGCTGCTG 4949  
Qy 4616 cccagaatgggtgagcagaatggccttgagatgagaggtcttgggttgggaagcagaag 4675  
Dh 4950 CCCAGATGGGTGACCAATATGCTGAGATGAGGATGAGGATGAGGATGAGGATGAGG 5009  
Qy 4676 acctctactgactcagcagatc 4699  
Dh 5010 ACCTTACTAGGGGTGGCCACTGC 5033

RESULT 11  
US-08-670-707A-38  
; Sequence 38, Application US/08670707A  
; Patent No. 5859204  
; GENERAL INFORMATION:  
; APPLICANT: Lollar, John S.  
; TITLE OF INVENTION: Hybrid Human/Animal Factor VIII  
; NUMBER OF SEQUENCES: 40  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.  
; STREET: 5370 Manhattan Circle Suite 201  
; CITY: Boulder  
; STATE: Colorado  
; COUNTRY: USA  
; ZIP: 80303  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/670,707A  
; FILING DATE: 26-JUN-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: WO PCT/US94/13200  
; FILING DATE: 15-NOV-1994



```

1      PRIOR APPLICATION DATA:
2      APPLICATION NUMBER: US 08/212,133
3      FILING DATE: 11-MAR-1994
4      PRIOR APPLICATION DATA:
5      APPLICATION NUMBER: US 07/864,004
6      FILING DATE: 07-APR-1992
7      ATTORNEY/AGENT INFORMATION:
8      NAME: Greenlee, Lorraine L.
9      REGISTRATION NUMBER: 27,894
10     REFERENCE/DOCKET NUMBER: 75-95F
11     TELECOMMUNICATION INFORMATION:
12     TELEPHONE: 303/499-8080
13     TELEFAX: 303/499-8089
14     INFORMATION FOR SEQ ID NO: 38:
15     SEQUENCE CHARACTERISTICS:
16     LENGTH: 4334 base pairs
17     TYPE: nucleic acid
18     STRANDEDNESS: double
19     TOPOLOGY: not relevant
20     MOLECULE TYPE: cDNA to mRNA
21     HYPOTHETICAL: NO
22     ORIGINAL SOURCE:
23     INDIVIDUAL ISOLATE: Factor VIII lacking B domain
24     FEATURE:
25     NAME/KEY: CDS
26     LOCATION: 3..4334
27     US-08-670-707A-38

```

Query Match	66.9%;	Score 3189.8;	DB 2;	Length 4334;
Best Local Similarity	83.5%;			
Matches 3692; Conservative	0;	Mismatches 637;	Indels 90;	Gaps 3;

[illegible]

QY	868	gaaggagctgcgcgaagaaagaaagaaagcctctgacaaatttatctacttcttgct	927
Db	603	gaaggagctctgaccagaaagaaagaaacccagaaacctgcagaaatttttactcttcttgct	662
QY	928	gtaatttgaatgaaggaaagttgtgacttcagaaacaaagaagcctctgtatgcagatagag	987
Db	663	gtcttttgattgaaggaaagaaagttggacttcagcaagaaatgacctcctggacacggccatg	722
QY	988	gattctgcatactgctcgagcctgtgcctaaatgcacaagctcaatggttatgtatgaacagg	1047
Db	723	gatcccgacacctgcgcagggccaccacctgcgaatgacacacactgaattgctgtcaacagg	782
QY	1048	tctctgcgaagttcgtattgagttgcgcgaagaaatcagttcatattgagatgtgattggaag	1107
Db	783	ttcttgcgcaagctctgattgagttgcgcgaagaaatcagttcatattgagatgtgattggaag	842
QY	1108	ggcacacctccctggaagtgcatacataattcctcgaagtgacacattctctgttgagagac	1167
Db	843	ggcacacctcccggaagtgacatactccattctcgaagtgacacattctctgttgagagac	902
QY	1168	catgcccagcgctccttggaaatctgcgcaataacttccctactgtctcctaaacactctg	1227
Db	903	catgccacgagcgtccttgcgaatgtccgccactgaatttctcctcactgtctcctaaacactctg	962
QY	1228	atggaaccttggaacagttctctactgtttttgttcatactctctccacaatagtatgycatg	1287
Db	963	atggacaccttggaacagttctctactgtttttgttcatactctctccacaataagtgatg	1022
QY	1288	gaagcttatgtcgaagtagaagctgtgcagagagaagcccaactaagaaatgaatgaat	1347
Db	1023	gaagcttgcacgtcagagtagaagctgtgcagagagaagcccaactaagaaatgaatgaat	1082
QY	1348	gaagaagcggaagactatgatgatcttactcagattcttgaatgtgatgtgtgacggtt	1407
Db	1083	gaagaagcggaagactatgatgatcttactcagattcttgaatgtgatgtgtgacggtt	1139
QY	1408	gatgatgaacactctcctctcttatccaaatctcgctcagttgtgcagaagacatccctaa	1467
Db	1140	gatgatgaacactctcctctcttatccaaatctcgctcagttgtgcagaagacatccctaa	1199
QY	1468	acttgggtacattcatgtctgtcgtgaagaggaagcgtggaactatgctcccttgctctc	1527
Db	1200	acttgggtgacattcatgtctgtcgtgaagaggaagcgtggaactatgctcccttgctctc	1255
QY	1528	gcccccgatgcacgaagttataaagctcaatatttgaacaatgcccccttaagcgtatgtc	1587
Db	1260	gcccccgatgcacgaagttataaagctcaatatttgaacaatgcccccttaagcgtatgtc	1319
QY	1588	aggaagatcaaaaaagttccgaatttatgtgcatacaacagatgaagacctttaagctcgtgaa	1647
Db	1320	aggaagatcaaaaaagttccgaatttatgtgcatacaacagatgaagacctttaagctcgtgaa	1379
QY	1648	gctatttcagacttgatcaagaaatcttggaagccttacttatatgggaagtttgagaacaa	1707
Db	1380	gctatttcagacttgatcaagaaatcttggaagccttacttatatgggaagtttgagaacaa	1438
QY	1708	ctgttgatcatatattaagaatcaagacaaagagacatatatacctcactcaacgaaatc	1767
Db	1440	ctgttgatcatatattaagaatcaagacaaagagacatatatacctcactcaacgaaatc	1499
QY	1768	actgatgtcgtctctctgttatttcaagagatataccaaaagtgttaaacaattgaagat	1827
Db	1500	actgatgtcgtcgtctctgttatttcaagagatataccaaaagtgtgtgaagaaagac	1559
QY	1828	tctccaattctgcgaaggaatatcttaaatataaatgtagcagttgactgtatgaagaatggg	1887
Db	1560	atggccaatttgcgacagagagaaacttttaaatataatgtagcagttgactgtatgaagaatggg	1619
QY	1888	ccaactaaataagaatccctcgtgtgcctgaacgcgtatactactcagtttgttataatggag	1947
Db	1620	ccaaccaaggtccgaatcctcgtgtgcctgaacgcgtatactactcagtttgttataatggag	1679
QY	1948	agaagctcactgattcaagactaatgtgcccctctctcatcgtctacaaagaatctgtatag	2007



Db	1660	AAAGATCTGGCTTGGGACATCATTTGGCCCTCTCTCATCTGCTACAAAGAACTGTGTAAAC	1739
QY	2008	caaaaggaagaaacccagaaataatgctcagaacaaggaatgcatccctgtttctcgtattatcag	2067
Db	1740	CAAAAGGAAACCAACGATGATGCTCAGACAAAGAAAGCTATCTCTGTTTCTGTATTGCAT	1799
QY	2068	gagaacccaagctgtgtactccacagaaatatacaacgcttctctcccaatccagctgtga	2127
Db	1800	GAGATCAAAAGCTGTGTACCTCCGACGAAATATTACACGCTTCTCTCCCAATCCGATGGA	1859
QY	2128	gtgcagctcttgagatcccaaggttccaaagctccaaatcaatgcaacgaatcaatggtctat	2187
Db	1860	TTACAGCCCCAGGATCCAGAGTTTCCAAGCTTCTAACATATCCACAGATCAATATGGCTAT	1919
QY	2188	gtttctgatatgttgacagttgtctcaagttgtttgcaagagtgagaaactgagtgacatctta	2247
Db	1920	GTTTTTGATAGCTTGGACAGCTGTGGGTTTGTTCACAGAGGTGGCTACTGATCAATTTCTA	1979
QY	2248	agcatlgagagacagactgtacttcctctctgtctctctctcgtgaatatactcttaaacac	2307
Db	1980	AGTGTTGGAGACAGAGAGGACTTTCCTCCGCTTCTCTCGGTACACCTTCAAAAC	2039
QY	2308	aaaatggtctatgaaagacacacatcccatlcccatctcgaagaaactgtctctcatg	2367
Db	2040	AAAATGCTATGTAGACACACACTCACCCCTGTTCCCTCTCAGGAAACGGCTTTCATG	2099
QY	2368	tcgatlgaaaaaacccagttcattgtattctcgtgggtgcacaaactgaactcttcgaagaga	2427
Db	2100	TCAATGGAAACCCAGGCTCTGTGGTCTTAGTGGTGCCAACTCAAGCTTGGGAAACGA	2159
QY	2428	gacatlgacgccttactcyaaggtcttctagtctgtgacaagaacacgtygatattatcagag	2487
Db	2160	GGGATGACAGCCTTACTGAAAGGTATAGTGTGACAGGGAATTTGGATTATTATATAC	2219
QY	2488	gacagttatgaaagatattcaagcatctgtcgtgagtaaaacaatgccaattgaaaccaaga	2547
Db	2220	AACACTTATGAAGTATTTCAGGCTTCTTGCTGAGGAAAGAAAGTCAATTGAACCCAGA	2279
QY	2548	agctctccagaaattlcaagaacacactagcactaggaagcaatltaatgcaacccca	2607
Db	2280	-----	2279
QY	2608	ccagttctlgaaacgccatcaacggagaataactcgtactactcttcaagtcagatcaagaag	2667
Db	2280	-----GACATTAAGCCTTCCTACTTTTCACCCGAGGAANAC	2315
QY	2668	gaaatgactatgatatgataccataatcaagttgaaatgaaagaaagatlttgaacattat	2727
Db	2316	AAAATGACATATATGATATCTTCTCACTCACTGAACCAAGGGAAGATTTTGACATTTAC	2375
QY	2728	gaatgaatgaaaatcagagccccccgacacttcaaaaagaaacacgacatatttat	2787
Db	2376	GCTAGGATTAATAATCAGACCTCTCCAGCTTTCAGAAAGAACCCGACACTATTTCATT	2435
QY	2788	gctgcagtggaagaggtcctbgya tla tggatgtagttagcttcccaacatgctctaaagaac	2847
Db	2436	GCTCGGTGGAGCAGCTCTGGGATTAAGGGAGTAGAGGATCCCCCGGGCCCTTAAGAAAC	2495
QY	2848	aggagctcagaatgagcagtgctccctcagtttcaagaaagtgttttccagaaattactgat	2907
Db	2496	AGGGCTCAGAACGAGAGGAGTGCCTCCGTTTCAAGAAAGTGTCCTCCGGAAATTTGCTAC	2555
QY	2908	ggctcccttactaagccctatbaccgttggaagaaactaaagaacatlttggtgcctcgtgag	2967
Db	2556	GGCCTCTTTCACGACGCGCTGTACCCGGGGAACTCACAAACACTTGTGGGCTCTTGGA	2615
QY	2968	ccatatataagagcagaagtgctgaaga taa latcatggttaacttccagaatccaggccct	3027
Db	2616	CCCTACATCAAGGGGAAGTTGAAGCAACATCATGTGTAACTTTCAAAAACAGCGCTCT	2675
QY	3028	cgctccattcttctattctagccttattctctatgagaaactcagaggaagcaagca	3087

Db	2676	GGTCCCAATTCCTTCCTACTGAGCCCTTATTCTTTATCCGGATGATCAGGACGCAAGGGGCA	27335
Qy	3088	gaacctagaanaaaacttgttcaagccctaatggaacccaanaacttacttggaaagtga	3147
Db	2736	GAACCTGCACACACACTTCGTGTCCACCCAAATGAAGAACGAACTTACTTTTGGAAAGTGCAG	2795
Qy	3148	catcatatgagcaaccactaaagaatgagtttgaactgnaaagcctgagctatttcttgat	3207
Db	2756	CATCACAATGGCACCACAGAAAGAGATTTGACTGCAGAAAGCCGGGCGCTACTTTCTGAT	2855
Qy	3208	gttgacctggaaaagaagtgtgacctgaagcgtgatttggaaccttctgtctgcacaact	3267
Db	2856	GTTCACCTGGAAAAAGATGTGCACTCCAGGCTTTATGTGGCCCCCTTCCTGTATCTGCGCGCC	2915
Qy	3268	aacacacgaacccctgtcctatgtagagaagaatgacagatcaagaatttgcctgtcttctc	3327
Db	2916	AACACCCCTGAACCGCTGCTCAACGGTAGACAGATGACCGTGCAGAAATTTCTGTGTTTTTC	2975
Qy	3328	accacatcttgaatgagaccaaaagcttgttacttccacttgaataatgynaagaacacagag	3387
Db	2976	ACTATTTTTTGATGAGACAAAGAGCTGTACTTCACTGAAAAATGTGGAAAGAAAGACTGCGCG	3033
Qy	3388	gtctccctgcaatatccagatgagaagatcccaactttaaagaagatatgcgttccatga	3447
Db	3036	GCCCCCTCCCACTCGAGATGTGAGAGACCCCACTCGTAAGAAAAAACTATCGCTTCCATGCA	3095
Qy	3448	atcaatgtgcataatgaatgataacacacccctgctgttagtaatgctcagaatcaaaagatt	3507
Db	3096	ATCAATGGCTTGTGATGATGATACACTCCCTGGCTTAGTATGTGCTAGATCAAAAGAGATC	3155
Qy	3508	cgatgtgtaactgcctcagcatgtaggaagaacatgaaacaatccatctatctatctacgttga	3567
Db	3156	CGATGTATCTGCTCACATGCGGAGCAATGAAGAAATATCCATTCGATTCATTTTAGCGGA	3215
Qy	3568	catgtgtcactctgtacgaaaaaagaagagatataaatggcactgtacaaatctctatcca	3627
Db	3216	CACGTTTCACATGTGACGAAAAAGAGAGATTAATAATGGCGGTGACAAATCTCTATTCGG	3275
Qy	3628	ggtgttttttggagaagtggaagaagttaaccatcccaagcctggaaatttggcggttgaattgc	3687
Db	3276	GGTGTCTTTGAGACAGTGGAAATGCTACCGTCCAAAGTTTGGAAATTTGGCGAATTAAGATTC	3333
Qy	3688	cttatltggcgaacatctcaatctgctgtggaatgagacacttltctgtgtacagcaataag	3747
Db	3336	CTGATTTGGCGAGACACTGCAAGCTGGGATGAGACAGCATTTCTCGGTGTACAGCAAGAG	3395
Qy	3748	tgtaagactccccctgggaatgagcttcttgacacattagagaatttccagatccagctcca	3807
Db	3396	TGTCAAGGCTCCACACTGGGAATGTGGCTTCTGGACGATTAGAGATTTTCCAGATCCACACTTCA	3455
Qy	3808	ggacaaatgtgaacagtgggcccccaaaagctggccagacttcatattccgagatcaatcaat	3867
Db	3456	GGACAGTATGGACAGTGGGCCCCCAAAAGCTGGCCAGACTTCATTTATTCGGATTCATTCAT	3515
Qy	3868	gcttggagacacaagagagcccttctcttgatcaatgaagtgtgactgttggcaccaatgata	3927
Db	3516	GCTGTGAGCACCAAGAGATCCCACTCTCTGATCAAGGTGGATCTGTGTGGCACCAATATGTC	3575
Qy	3928	atcaacggcatcaagaaccacagagtgccgttcagaagttctccagcctcaatccatctccag	3987
Db	3576	ATTTACACGGCATTCAGACCCAGGGGTGCCGCTCAACAAATTTTCCAGCCTCTACATCTCCAG	3633
Qy	3988	tttctacatcagtatagttcttgttggagaagaatgtagcacacttaccggagaatctcaact	4047
Db	3636	TTTATCATCATGTACAGTCTTTGACGGGAGGAACCTGGCAAGTTTACCGAGGGAATTCACG	3695
Qy	4048	ggaaccttaatgagctctcttcttggcaatgtgatatctccggagataaaacacatatcttct	4107
Db	3696	GGCACTTAATAGGTCCTCTTTGGCAATGTGAGACGATCTGTGGATTAAACAAATATTTTTT	3755
Qy	4108	aacctccaatlatgtctgatacatccgttcttgaccaaccaactatataagcatcgcagc	4167
Db	3756	AACCTTCGATTTGTGGCTGCGGTACATCCGTTTGACCCAAACATTTAAGCATTCGCGAGC	3815



QY	4168	acccttcgcaatggagtttgatggagctgtgatttaataagtttgcagcatgcattgggaatg	4227
Db	3816	ACTCTTGGCATGGAGTTGATGGGCTGTGATTTAAACAGTTGCACATAGCCCCCTGGGAATG	3875
QY	4228	gagagtaagaacaaatcatcagaatgcacagattactgtcttcatctactcttacaatattgt	4287
Db	3876	CAGAAATAAAGCGATATCATGACTGCACAGATCATCGGCGCTCTCCACGTAAGCAATATATTTT	3935
QY	4288	gccacctgtgtctcttccttcaaaaagcttcgaactctcaactctcaaggaggagtaatgtctgaga	4347
Db	3936	GGCACCCTGGCTCTCTTCCACAAAGCCCGACTTCACCTCCAGGGGCGGACGAATAGCTGGCGA	3995
QY	4348	cctcaggttgataataatccaaagaagtgctgtcgaatgagacttccagaagacatatgaatc	4407
Db	3996	CCCCGGGTAGACGAGCGCAGAGAGATGGCTCTCAGTGTGACCTGCAGAGACCGGTAAAGTC	4055
QY	4408	acagaagtaactactctcaaggagtaaaatctctctgtcttaaccagcatgtatgtgaagagttc	4467
Db	4056	ACAGGCATCCACCCACGAGGCGGTGAATCCCTGCTCAGCAGCATGTATGTGAAGAGATTG	4115
QY	4468	ctcatctccagcagcagaagaatggccatcagtgagactctcttttttcagaaatggaaagta	4527
Db	4116	CTCGTGTCCAGTGTGATCAGGACGGCCGCCCTGACCTGTGTTCTTCCAGGACGGCCACACG	4175
QY	4528	aaggttttccagaagaaatcaagaactctctcacacctgtgtgaaactctctagaaccacog	4587
Db	4176	AAGGTTTTTCAGGGGCAATCAGAGACTCCTCCACCCCTGGGAGACGCTTGGAGACCCCG	4235
QY	4588	ttactgactcgcactcttgaaatcccccagagttgggtgtgaccacagattgacctag	4647
Db	4236	CTGTTACAGCGGCTACCTCGAAGATCCACCCACGACGCTGGGGGCGACATGCCCTGAGG	4295
QY	4648	atggaagttctggcgtgcagagacacagagactctactga	4686
Db	4296	CTCGAGGTTCTAGAGATGTGAGGCACAGAGATCTGTACTGA	4334
RESULT 12			
US-09-037-601-38			
; Sequence 38, Application US/09037601			
; Patent No. 6180371			
; GENERAL INFORMATION:			
; APPLICANT: Ioliar, John S.			
; TITLE OF INVENTION: Hybrid Human/Animal Factor VIII			
; NUMBER OF SEQUENCES: 40			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.			
; STREET: 5370 Manhattan Circle Suite 201			
; CITY: Boulder			
; STATE: Colorado			
; COUNTRY: USA			
; ZIP: 80503			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Floppy disk			
; COMPUTER: IBM PC compatible			
; OPERATING SYSTEM: PC-DOS/MS-DOS			
; SOFTWARE: Patentin Release #1.0, Version #1.30			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/09/037, 601			
; FILING DATE: 26-JUN-1996			
; CLASSIFICATION:			
; PRIOR APPLICATION DATA:			
; APPLICATION NUMBER: WO PCT/US94/13200			
; FILING DATE: 15-NOV-1994			
; PRIOR APPLICATION DATA:			
; APPLICATION NUMBER: US 08/212,133			
; FILING DATE: 11-MAR-1994			
; PRIOR APPLICATION DATA:			
; APPLICATION NUMBER: US 07/864,004			
; FILING DATE: 07-APR-1992			
; ATTORNEY/AGENT INFORMATION:			
; NAME: Ferber, Donna M.			

```

?      REGISTRATION NUMBER: 33,878
?      REFERENCE/DOCKET NUMBER: 75-95F
?      TELECOMMUNICATION INFORMATION:
?      TELEPHONE: 303/499-8080
?      TELEFAX: 303/499-8089
?      INFORMATION FOR SEQ ID NO: 38:
?      SEQUENCE CHARACTERISTICS:
?          LENGTH: 4334 base pairs
?          TYPE: nucleic acid
?          STRANDEDNESS: double
?          TOPOLOGY: not relevant
?      MOLECULE TYPE: cDNA to mRNA
?      HYPOTHETICAL: NO
?      ORIGINAL SOURCE:
?      INDIVIDUAL ISOLATE: Factor VIII lacking B domain
?      FEATURE:
?      NAME/KEY: CDS
?      LOCATION: 3..4334
?      US-09-037-601-38

```

Query Match	Similarity	66.9%	Score 3189.8	DB 4	Length 4334
Best Local	Similarity	83.5%	Pred. No. 0		
Matches	3692	Conservative	0	Mismatches 637	Indels 90
					Gaps 3
QY	271	atcgcaatagagctctccacgcgtctcttccttcgctgctcttcgcatlccgcatltagtgc	330		
Db	3	ATGCAGGCTAGAGCTCTCCACCTGTGCTCTTCTGCTGCTCTTCCCACTCGCGCTTTAGTGGC	62		
OY	331	accgaagaatctactacacctgggtgacgtgaaatgtaacatgggactatatacgaaagtga---t	387		
Db	63	ATCAGAGATATACCTACTCGGCGCCAGTGAACATGCTCTGGGACTACCGGCAAGTGAATC	122		
OY	388	ctcagtgagctgctccgtgtagcagcaagattctctctctagagtgccaaatcttccatlc	447		
Db	123	CTCCCTGAGCTGCACGTGGACACACGATTTCTCGTTACAGCCAGAGACTTCTCCGTTG	182		
OY	448	aacctcgaatcgtgtacaaaagaagctcgtgtgtgagaatccagaagttcaacttccaac	507		
Db	183	GGCCCGTCACTCTTACAAAAGACTGTGTTCTGTAAGTTACGGGATCACTTTTCAGC	242		
OY	508	atcgctaaagccaaagccacccctgtagtggtctgctaggttccctacacatcaggtcgaagtt	567		
Db	243	GTTGGCAGGCCCCAGGCCACACATGATGGGTGTGCTGGGGTCTCATCAGCTCAGGCTAGGTT	302		
OY	568	tatgatacagtggtgctaatcacttaagaacatlgcttccacatcctgtaagttcgaagct	627		
Db	303	TACGACAGGCTGTGGTTGTTACCTCTGAAGAAACATGAGCTTCTCATCCGTTTAGTTCACGCT	362		
OY	628	gttgtgtatcctacttggaagctcttcgtgagagcttgaaatagatagatcaagccagctaa	687		
Db	363	GTCCGCGTCTCTCTTGTGGAATCTTCCGGAAGCGCTGAATATGAGAGTACACAGCCAA	422		
OY	688	agggagaagaagaatgataaagctctcccttggtggaagccatacatatgctgcgaagtc	747		
Db	423	AGGGAGAAAGGAAGACGATTAAGTCTCTCCCGGTTAAAGCCAAACTAGCTCTGGCAGGGTC	482		
OY	748	ctgaagaaggaatggtccaaatgagcctctgacccactgtagcttaccctactacatatactct	807		
Db	483	CTGAAGAAGAAATGTCACCAACACGCTGTAGCCACCATGTCTACCTACTACATACCTGTCT	542		
OY	808	catgtgagacctgtaaaagacttgaattcaagcctcaattgagagccctactglatgtataga	867		
Db	543	CACGTGAGACTGTGTAAGAAAGACTGGAATTCGGGCTCATATTGGAGCCCTGCTGTTGTGAGA	602		
OY	868	gaaggagatctggtccaaaggaagaagacagacacctgacacaattataactacttlttgt	927		
Db	603	GAAAGGAGCTGTGACCAAGAAAGGCCAGAACCTGCAAGAACTTGTGTACTCTTTTGTCT	662		
OY	928	gtatttgatgaaaggaagaagtgtgacatcagaacaagaagactccttgatgagaatgag	987		
Db	663	GCTTTTGATGAAGGAGAAAGTGTGGACTCAGCAAGAAATATGACTCTCTGGACACAGGSGCATG	722		



[illegible]

Db	1800	GGAAATCAAAAGCTGGTACTCTGCGAGAAATATTCAGCGCTTCTCCCAATCCGATGGA	1859
Qy	2128	gfgcagctctggagatccagagttccagagctcccaacatcagacaagcaaaatgagcat	2187
Db	1860	TTACAGGCCCCAGGANTCCAGATTCCTCAAGCTTCTACATCATGACAGCATCAATGGCTAT	1919
Qy	2188	gtttctgtagtcttcgagttgtcagttgttgttcagtatgagtggtgacatcgttactcta	2247
Db	1920	GTTTTTGATACCTTTCGACGCTGTGGTTGGTTTTCACAGAGGTGGCATCTGTATCTCTA	1979
Qy	2248	agcatttgagacacagactgaacttccttctcgtctcttctctctcgtgatataccttaaacac	2307
Db	1980	AGTGTGGAGACAGACGCGAGCTTCCTCCGTCTCTCTCTGCGTACACCTTTCAAACAC	2039
Qy	2308	aaaatgycctcttgagagacacatcacctctatcccatcttcagtgagaacatcttctcatg	2367
Db	2040	AAAATGGCTATGGAAGACACACTACCTGTCTCCCTTCTCAGAGAAACGGTCTTTCATG	2099
Qy	2368	tcgatygaaaaacccaagtlctatgtagtctctgggtgtgcacaactcagactcttcgaacaga	2427
Db	2100	TCAATGGAAAACCCAGGTCCTGTGGGTCCTTAGGGTGCCAACTCAACACTTTGGGAAACAGA	2159
Qy	2428	ggcattgacgcgcttacttgaagagtttctagtctgtgacagaagaacatcgtgtatattacgag	2487
Db	2160	GGGAAGACAGCCTTACGTGAAGGTGTATAGTGTGACAGGAGCATTGGCTATTATATGAC	2219
Qy	2488	gacagtttgagaagatattcttcagcaactctctctgtagtcaaaaacatccatctgaaccaga	2547
Db	2220	AACACTTATGAAAGATATTTCCAGGCTTCTTCTCTAGTGGAAAGATGTCAATTAAACCAGA	2279
Qy	2548	agctctctccagaatltcaagaacacctagcactaggcaaaagcaattlaatgtcacacca	2607
Db	2280	-----	2279
Qy	2608	ccagtccttgaaacgccaatcaacgggaaataactcgtactactcttcagtcagatacgaag	2667
Db	2280	-----GACATTAAGCCTTCTCACTTTCACGCGGAGGAAAC	2315
Qy	2668	gaatttgactatgtataccatatcagtttgaaatgaagaagaagatttgacattat	2727
Db	2316	AAAATGGACTTGTGATATTCCTTCTCACTGAAACGAGGAGAAATTTTATCATTTTAC	2375
Qy	2728	galtgagatcgaaaaatcagagcccccagcttctcaaaagaanaaacagacactatttlat	2787
Db	2376	GGTAGAGATGAAATTCAGAGACCTCGGAGCTTTCAGAAAGAGAAACCGAATATTTCAT	2435
Qy	2788	gctgcagtgagagagctctcggattatlyggaatgagtagtctcccaatgltctlaaanaac	2847
Db	2436	GCTGCGGTGGAGAGCTCTGGGATATTCAGGATTCAGCGAATCCCCCGGCGCTTAAGAAAC	2495
Qy	2848	agggtctagagtggtgcagtgctgccttcagttcagaagaagtgtgtttccaggaattactgat	2907
Db	2496	AGGGCTTCAGAACGGAGAGTGGCTCGGTTTCMAAAGTGGTCTCTCCGGAAATTTCTCAC	2555
Qy	2908	ggtcccttactcagcccttatatacgttlygagaactaaatgaaacatttggagctccctggg	2967
Db	2556	GGCTCTTTCACAGCGCGGTGTACCGGGGGAATCAAAACACTTGGGCGCTCTTGGGA	2615
Qy	2968	ccatatataagagcagaagttgaaagataatataatcagtgtaacttcaagaatcagagctct	3027
Db	2616	CCCTACATCAAGAGCGGAAGTTGAAAGAACATCATGTGTAACCTTTCAAAACACGAGCGTCT	2679
Qy	3028	cgctccattcttcctctatcttcagacttatcttcttaagagaagaatcagagaggaagaga	3087
Db	2676	CGTCTCTATTCCTTCTACTGAGCTTATTTCTTATTCGGAATGATGAGGACGAAGGGCA	2735
Qy	3088	gaacctagaaaaaacttgtcgaagccttaatlgaaacccaacacttacttctggaagtgcac	3147
Db	2736	GAACTGACACACACTTCTGTCAGGCCAAATGAAGAACAGAACTTACTTTTGGAAATGCGAG	2795
Qy	3148	calcatatgacccactaaagaatgagtttgactgcaaaagccttggctattattctctgat	3207



Db 2796 CATACATGACACCCAGAGACGATTTGACTGCAAAAGCCTGGCCCTACTTTTCTGAT 2855  
Qy 3208 gttgacctggaanaagatgtgcaactcaagcctgatttgagccctcttggtctgcacact 3267  
Db 2856 GTTGACCTGGAAAAAGATGTGTGACCTGAGCTTGATGGCCCTTGTGATCTGCCGCGC 2915  
Qy 3268 aacacacgtgaacctgtctcaatggagacaagtgcacagtaacaggaatttgcctgttttc 3327  
Db 2916 AACACCCGTGAACGCTGCACGCTGACAGTAACTGACCGTCAAGAAATTTCTGTTTTC 2975  
Qy 3328 acctctttgtatgagaccaaagctgtgtactctacatgaaatatggaagaacctcaag 3387  
Db 2976 ACTATTTTGTAGAGACAAAGAGTGTGATCTTCACTGAAAATGTGAAAAGAACTCCGG 3035  
Qy 3388 gtcctctgcaatatccagatgtgaagaatcccaactttaaagaagaatatactgtctcaatga 3447  
Db 3036 GCCCCTCCCAACCTGCAGATGAGAGCCCACTGTGAAGAAACATATGCTTCATGGA 3095  
Qy 3448 atcaatgtctacataatgtgatacactacctgtgcttaataatgtctcaagaatcaagaatt 3507  
Db 3096 ATCAATGGCTATGTGATGGATACACTCCCTGGCTTAGTAATGGCTCAGAAATCAAGATC 3155  
Qy 3508 cgaatgtatctgtctagatgtggcagcaatgaataatccatctctatctatctagtgga 3567  
Db 3156 CGATGTGATCTGTCTGATGGGAGCAATGAAATATTCATTTCGATTTTACGGGA 3215  
Qy 3568 catgtgtcactgtacgaaaaaagaagagataaaatgtgcaactgtacatctctatca 3627  
Db 3216 CACGTGTTCACTGTACGGAAAAAGAGATGTAATAATGCCGTGTACAAATCTCTATCCG 3275  
Qy 3628 ggtgttttttgagacagtggaatgttacaatccaaagcttgaaatttgcggtgtggaatgc 3687  
Db 3276 GGTGTCTTTGAGACAGTGGAAATGCTACCGTCAAAAGTTGGAATTTGGCAATAGAAATGC 3335  
Qy 3688 cttaattggcagagactctatcaatgttgatgtgagacacacttctctgtgtgaaggaataag 3747  
Db 3336 CTGATGTGGAGACCTCTCAAGCTGGATGAGACACACTTCTCTGTGTACAGCAAGAG 3395  
Qy 3748 tgtcagaactcccttggaatggctctctgacacatlaagaattctcaagaattcaagcttca 3807  
Db 3396 TGTGAGGCTTCACCTGGAAATGGCTTCTGAGCAGCATTAGAATTTTCAGATCACAGCTTCA 3455  
Qy 3808 ggaacaatgtgaacagtggtgccccaaagctgtgcgaactctatctcgatcaatcaat 3867  
Db 3456 GGACAGTATGACATGGGCCCCCAAGCTGGCCAGACATTCATTATTCGCGATCAATCAAT 3515  
Qy 3868 ggcctggaacaccaaagagcccttctctggaatcaaggttgatctgtgtgacacaaatgatt 3927  
Db 3516 GCCTGGAGACCAAGATGCCCACTCTGTGATCAAGTGTGATGTGGACCAATGATGC 3575  
Qy 3928 attcaaggcatcaagaacccagggtgcctgtcagaagttctcagacctctacatctctcag 3987  
Db 3576 ATTCAAGGCATCATGAGCCAGGCTGCCGTGAGAAATTTTCCAGGCTCTACATCTGCCAG 3635  
Qy 3988 ttatcatcatgtatgtctctgtgaatggagaagtggtgagaactctatctcgaagaattccact 4047  
Db 3636 TTTATCATCATGTACAGTCTTGACGAGGAGAACTGGCAGAGTACCGAAGGAAATTCACAG 3695  
Qy 4048 ggaaccttaaatgtctcttttggaatgtgaattcatctcggataaataacaaatctttt 4107  
Db 3696 GGCACCTTAATGTTCTTTTGGCAATGTGGAGCATCTGGGATTTAAACACATATATTTT 3755  
Qy 4108 aacctccaatatattgtctcgaatacgtttgtcaccccaaccatcatatagcatctgcagc 4167  
Db 3756 AACCTCGATTTGTGGCTCGGATCACTCGTTTGCACCCAAACATTTACAGCATCCGCAAC 3815  
Qy 4168 acctctgcagatgtgaatgggtgtgacttaataatgtctcagatcgcatttgggaatg 4227  
Db 3816 ACTCTTGCAATGAGATGATGGGTGTGATTTAAACAGTTGCAGCATGCCCCGGGAATG 3875  
Qy 4228 gagaagtaagaataatcagatgcagaaattactgtctcatctcatcttaacatatgttt 4287  
Db 3876 CAGAAATTAAGGATATACAGACTCACAGATCAAGGCTCTCTCCACCTAAGCAATATATTT 3935

Qy 4288 gccacctgtctctctcaaaaagctcgacttcaactctccaaaggaggaatgatccttgga 4347  
Db 3936 GCCACCTGGTCTCTTACACAGCCGCACTTCACTCCAGGGCGGACGAATGCTCGCGGA 3995  
Qy 4348 cctcagtgtaataatccaaaagatgtgctgcgaatgtgaccttccagaagaatgaatgc 4407  
Db 3996 CCCCCGTGAGAGCCCAAGAGATGTGCTGACAGTGTGACCTTCAACAAACGGTGAAGTTC 4055  
Qy 4408 acagagtaactactcaggaggaataatctctgtaccagatgtatgtgaagagctc 4467  
Db 4056 ACAGGATACCAACCCAGGCGCTGAAGTCTCTGCTCAGCAGCATATATGTGAAGAGTTC 4115  
Qy 4468 ctcaatccagaagtaatgaatgtgcatacagtgactctctcttcttcaagaatgtcaagta 4527  
Db 4116 CTCGTGTCCAGTATGAGAGAGCGCGCCCTGACCCCTGTTCCTTCAGAGCGGCACAG 4175  
Qy 4528 aaggttttccagggaatcaagaactctctacacactgtgtgtgaactctctcagaaccag 4587  
Db 4176 AAGGTTTTCAAGGGCAATCAGGACTCTCCACCCTGTGTGAAGCGTCTGGACCCCGC 4235  
Qy 4588 ttactgactcgtactccttgaaattcaaccccaagattgtgtgaccagatgtgccttagg 4647  
Db 4236 CTGTTACGCGCTACTCTGAGATCCACCCACGACTGGGCGACACATGCGCTGAGG 4295  
Qy 4648 atggaagttctgtgctgcagagcacagagacctctactga 4686  
Db 4296 CTCGAGGTTCTAGATGTGAGGACACAGATCTCTACTGA 4334

RESULT 13  
US-08-717-294-42  
; Sequence 42, Application US/08717294  
; Patent No. 6114148  
; GENERAL INFORMATION:  
; APPLICANT: SEED, BRIAN  
; APPLICANT: HAAS, JURGEN  
; TITLE OF INVENTION: HIGH LEVEL EXPRESSION OF  
; TITLE OF INVENTION: PROTEINS  
; NUMBER OF SEQUENCES: 110  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Clark & Elbing LLP  
; STREET: 176 Federal Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02110  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/717,294  
; FILING DATE: 20-SEP-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Elbing, Karen L  
; REGISTRATION NUMBER: 35,238  
; REFERENCE/DOCKET NUMBER: 00786/345001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-428-0200  
; TELEFAX: 617-428-7045  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 42:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4451 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear



Page 31

US-08-717-294-42

Db 1044 GAGGCTTTACGTGAAGGTGGACAGCTGGCCCCGAGAGAGCCCCAGCTGGCGCATGAAGACAC 11033

Qy 1348 gaagaagcgaagaactagatgatatctactgattcgaatgatatgtgcacgitt 1407

Db 1104 GAGGAGGCGCGAGGACTACGACGAGACCTTACCGACGAGAGATGGATGTCTCTACGCTTC 1163

Db	1044	GAGGCTTACGTGAAGGTGGACAGCTGCCCCCGAAGACCCCACTGGCATGGAAGAACAC	1103
Qy	1348	gaagaagggagaagactatgatgatgacttaacgacttccgaatgagtgtgacggtt	1407
Db	1104	GAGGAGGCGCCGAGGACTACGACGACGACCTTACCAGACGAGAGTGGATGTCTACGCTTC	1163
Qy	1408	gatgatgaacaatctccctcccttatccaaatctgctcagtttgcaagaagaatccctaa	1467
Db	1164	GACGACGACACACGCCCGAGTTTCATTCAGATCCGAGGGTGGCCAAAGAACGACCTTAAG	1222
Qy	1468	acttgggtacatctacatctgctgtctgaagaggagacttggagactatgctccctagctctc	1527
Db	1224	ACCTGGGGGCGACTACATCGCGCGGAGAGAGAGAGAGTGGGACTACGCCCCCGCTACTACTG	1283
Qy	1528	gcccccgatgcagaagtcttaaaatgcatttjaaaanaatgagcccttaagggatgtgt	1587
Db	1284	GCCCCCGACGACCCGACTCTCAAGAGCCACTACTTACACAGGCCCCCAGCGCATCGCG	1343
Qy	1588	aggaagtcaaaaaagtcgcgatttatgcaacacagatgaaaccttlaagaactctgtgtaa	1647
Db	1344	CGCAAGTACAGAAGGTGGCGTTTCATGGCTTACCGGAGAGACTTTCAAGACCGCGAG	1403
Qy	1648	gctattcaagatgaatcaaggaatcttgggaacttacttlatggygaagtgtgagaaca	1707
Db	1404	GCCATCCGCGCAGAGTCCGGCATCTCTGGCCCCCTCTGTACGGCGAGGTGGGCGCAACCC	1463
Qy	1708	ctgttgatataatttaagaatcaagaagaagacatataaatctacccctacaggaatc	1767
Db	1464	CTGTCTGATCATCTTCAAGAACCAAGGCGAGGCGCTTACACACTTCAACCCCAACGGCATC	1523
Qy	1768	actgatgtcgcttcccttggtlatccaagagatacccaaaagtgyltaaaacatttgaagat	1827
Db	1524	ACCGACGTGGCCCCCTGTACAGCGCGCGCTGCCCAAGGGGTGAAGCACTGAAGGAC	1583
Qy	1828	tttccaatctgcagaagagaataatccaataaataatgacagtgactgttagaagatgg	1887
Db	1584	TTCCCCATTCCTGCCCCGCGGAGATCTTCAAGTACAACTGAGACCGTGAAGGACGCGC	1643
Qy	1888	ccaactaatcagttctctggtgcgcgacccggtatctactagtgttgcgttaatatgag	1947
Db	1644	CCACCAAGACCGACCCCGCTGCTTACCCGCTTACAGCAAGCTTCTGTGAACATGGAG	1703
Qy	1948	agaagatcagtttcaagcactatgtggccctctctcatctgctatacaaaagactctagat	2007
Db	1704	CGGCACTCGGCGCTCCGGAAGTATCGGCCGCCCTGCTATCTGTACAGAGAGAGCGTGGAC	1763
Qy	2008	caaaagagaaacccagataatgtcagaagaagaaatgtcalccgttltctgtlatltgat	2067
Db	1764	CAGGCGCGGACACAGATCATGAGCGACAAAGCCAAAGTATCTGTACGCTGTTCGAC	1823
Qy	2068	gagaacogaagcttggttaccctccacagagaaatatacaacglttctcccaatccagcttga	2127
Db	1824	GAGAACCGCACTGTGTCTGTACCGGAATAATCAACGCTTCTCTCCCAACCCCGCTGGC	1883
Qy	2128	gtcgaagcttggagatcccaagattccaagcctccaaatcatgtcaagcaatcagtgcat	2187
Db	1884	GTGCAAGCTGGAAGATCCCGAAGTTCACGACGAGCAAAATATGATCAGACATCAACGGCTAC	1943
Qy	2188	gtttttgatatgttgcagttgtcagtttgttttgatgatgagtgacatcactgtatctcta	2247
Db	1944	GTTGTTCCGACACCTCGTAGCTGAGCGTGTGCTCATGAGAGTGGCTACGTGATCAATCTCG	2003
Qy	2248	agcatttggagacagagactgactctcttctgtcttcttctctcgtgataactcttaaaac	2307
Db	2004	AGCATTCGGCGCCCGACCGACTTCTGTAGCGTGTCTTCTCCGGGTATTAACCTCAAGAC	2063
Qy	2308	aaatgtgctatatgaaagacacactcacocctatccactctcagaggaagactgtctctag	2367
Db	2064	AAGATGGTACGAGGACACCTTGACCTGTTCCTTCTCCGGGAGACATGTGTTCATG	2123
Qy	2368	tcgatgtgaaaaaccaggtctcatgtatcttctgtggtgtgcacaacatcagacttcttgagaca	2427



Db	2124	TCATATGGAGAACCCCGGCGCTGTGATCTTGTGGCTGCCAACAAACACGACCTTCGCAACCGC	2185
Qy	2428	ggcactgacgcgccttactgtaagagtttccagttctgtgacaaagacaclygtatgattacgag	2487
Db	2184	GGCATGACTCCCTGCTGTAAGTCTCCAGCTGCACAAAGAACCCGGGAGTACTACGAG	2243
Qy	2488	gacagttatgaagatatattcaagcaacttctgtagtataaaacaaatgcatctgtaacaaaga	2547
Db	2244	GACAGGTACGAGACATCTCCGCTTACTCTCTGTCCAAAGAACAGCGCATGTGACCCCGC	2303
Qy	2548	agcttctcccaagatccaagacacccctcagcaactgagcaagaagcaattaaagtcaacccca	2607
Db	2304	TCTCTTCCCAAAACTCCCGCCACCCGACGACGGGTGAGAAGGAGTTCAAGCCACACCCC	2363
Qy	2608	ccagttctgaaacgcacatcaacgcgggaaataaactgatactacttcttcagttcagataaag	2667
Db	2364	CCCGTCTGTGAAGGCGCCACAGCGGAGATCACCCGCGAACCCCTGTCAAGAGGACAGGAG	2423
Qy	2668	gaaattgacatatgtatataccataatcaattgtgaaatggaagaagaaatttgaacttatt	2727
Db	2424	GAGATGTACTACACGACACCATATCAGCGTGGAGATGAAGAGGAGACTTGACATCTAC	2483
Qy	2728	gatgtagaatgtaaaatacgaagccccgcagctttccaagaagaacagcaattatatt	2787
Db	2484	GACGAGGACGAGAACCCAGACCCCGCTCTCTCCAAAAGAAAACCCGCACTTCACTCATC	2543
Qy	2788	gctbcagctggaagagcctctggaattatggaatggaatgactccacatlytcttaagaaac	2847
Db	2544	GCGCGCTGTGAGCGCTGTGGGACTACGGCATGAGCAGACCCCGCATGCTCTCGCAC	2603
Qy	2848	agaggtccagagttgcaagtgtccctcagtttcaagaagattgtttccaagaaattacgatt	2907
Db	2604	CGCGCCAGAGCGCGACGCGCCCACTTCACAAAGAGTGGTTCAGAGACTTCACGAC	2663
Qy	2908	ggctccttcaactcagccctatataccgctggaagaaatgaacaattggaactcctgag	2967
Db	2664	GGGAGTTCAACCCAGCCCGCTGTACCGGGGGAGCTGAACAGACACTGTGGCTGCTCGCG	2723
Qy	2968	ccaataataagacgaagaagttgaagaataataatcagtgtaacttcagaatcagagcctct	3027
Db	2724	CCCTACATCCGCCCGGAGTGAGGACMACATCATGTGTGACTTCCGCAACCAAGCCCTCC	2783
Qy	3028	gctcccatatcccttcaattctcagcttatattcttatagaagaagaatcgaagcaagaagca	3087
Db	2784	CGGCTTACTCTTCTTACTCTCTCCCTATCAGCTACAGAGGAGACCAAGCCGACGGCGCC	2843
Qy	3088	gaacactagaanaaacttctgcaagcccaatgaaaccaaacttacttcttggaagaagtcaa	3147
Db	2844	GAGCCCGGACGAAGACTTGTGAAAGCCCAAGACGACATTAAGACTTCTGTGAAGGTGAG	2903
Qy	3148	catcataatggaccacacataagaatgattgtgaactgcaaaagcctgaggttattctctgat	3207
Db	2904	CACCATATGGCCCCCAACCAAGGACGATTTGTGACGTCAAGGCTGTGGCTTACTTACGGCAC	2963
Qy	3208	glttgacctggaanaaataatgtgcactcaagagctgaattggaaccccttctggtctgcaact	3267
Db	2964	GTGGACTGTGAGAGGACGTGCACAGCGGCTGTATCGCCCCCTGTGTGTGTGCACACCC	3023
Qy	3268	aacacacitgaacccttgcataatggagaagaatgagatcacagaatttgcctgttcttc	3327
Db	3024	AACACCTGAACCCCGCCCAACGSGAGGAGTGACTGTGCAGGAATTTGCCCTCTTCTTC	3083
Qy	3328	accacatcttgaatgaaccaaagcctgttactcactcgtgaaataatgaaagaanaactgcag	3387
Db	3084	ACCATCTTGTGACGAGACTTAAGAGCTGTACTTCCACCGAAGCAATGTGAGCGCACTGTCCGC	3143
Qy	3388	ggtccctgaatatctcaagatgtaagaatcccaactttaagaagaatttgcgttccatgca	3447
Db	3144	GCCCTCTGACACTCCAGATGGAAGATCCCACTTCAAGGAGAACTTCCGCTTCACGCC	3203
Qy	3448	atcaatcgtctataataatgataaactcctggtcttaataatggtctcgaatcaagaat	3507
Db	3204	ATCAACGGGTACATCATGTGACACCCCTCGCGCTGTGTGATGTGCCACGAGACAGCGATC	3263

QY	3508	cgaatgatactctgctcagcatatggcagcaatgaaacatccattctatcttaattcgaatgga	3566
Db	3264	GCCTGGTACCTGCTGCTATATGGGACGACAGAGAACATCCACAGCATTCACCTTCCAGCGGC	3323
QY	3568	catgtgttctactctgtacgaaaaaagaagagatataaaatggcaatgtacatactcatca	3627
Db	3324	CACGTTTTCACCGTGGCCGAGAAAGAGAGAGTACAAAGATGGCCTGTACAACTGTACCCC	3383
QY	3628	ggtgttttttggaaagtggaaatgttaccatccaaagtggaaatttggcgggtggaaatgc	3687
Db	3384	GGCGGTTCGAGACTGTGGGAATGCTCCACGAAAGCGCGGATCTGGCGCGTGGAGTGC	3443
QY	3688	cttatgtgcagagcatctacatgctcgggatagagaacttttctgtgtgtacagaaatga	3747
Db	3444	CTGATCGCGCGAGACCTTGCAAGCGGAGTATGACCTGTTCTGTGTATGACGAAACAG	3503
QY	3748	tgtcagactctcccttgggaatggtctcttcgcacacatlagagatlttcaagatlaacgttca	3807
Db	3504	TGCCAGACCCCCCTTGGGATGGCCAGCGGCGACATCCGAGACTTCCAGATTCACCCGACG	3566
QY	3808	ggacacataatggacagatggggccccaagaatctggccaagacttcatatttcggatcaatcaat	3866
Db	3564	GGCCAGTACGGCGAGTGGGCTCCCAACCTTGCCGCTGACACTACAGCGGAGCATCAAC	3623
QY	3868	gcttggagcaccaagagagcccttctcttggatcaagttgatatgtgtgtggcaacatgatt	3927
Db	3624	GCTGTGTGACCAAGAGAGCCCTTCTCTGGATGACAGGTGACCTGCTGGCCCCCATGATC	3683
QY	3928	attcaagcatcaaaagaccagaaggtgtccgttcagaagltctcgaactctacatctccag	3987
Db	3684	ATCCACGGCATACAGACCCAGGGGGCCCGCAGAAAGTTACAGACCTGTACATCAGCCAG	3743
QY	3988	tttatacatcatgataagtcattgataaggaaagttggcagaactatctcgaaggaaattccact	4047
Db	3744	TTTCATCATCATGTACTCTCTTAGAGCGGAAGATGGCAGACCTACCCGGCGCAACGACCC	3803
QY	4048	ggaaccttaatgtgctctctcttcttggcaatgttgattatcttgggataaaacaataatttt	4107
Db	3804	GGCAGACCTGATGGTGTCTTGTGGCAAGTGAAGACAGCGGCTACAGCAACATCTTTC	3863
QY	4108	aaacctccaattatgtctgataataacgtttgcacccaactcatatgatagattgtgcagc	4166
Db	3864	AACCCCCCATCATGTGCCCCGTACATCGCGCTGCAACCCACCATCTACAGCATCCGACGC	3923
QY	4168	actcttcgcatggagatttgatgtgctgtgatttaaatagttgcagcatggcatttggaaatg	4227
Db	3924	ACCCGGCCATGGAGCTGATGGCTGGCCTGCACTGAACGCTGACGATGCCCTCGGCGATG	3983
QY	4228	gagagtaaaagcaataatacagatgtgcacagatattactgtcttcatctactcttacaatatgttt	4287
Db	3984	GAGACCAAGGCGATACAGAGAGCCCAATACCGCTCCAGCTACTTACCAACAATGTTTC	4043
QY	4288	gcaacctgtctctcttcaaaagctgcgacttcacctccaaggagagagtaatgcctggagaa	4347
Db	4044	GCGACTCTGGAGCCCGCAGAGAGGCCCGCTTACCTGTGAGGGCGCGCAACGCCCTTGGCG	4103
QY	4348	cctcagatgtaataatccaaaagatgtgtcgaatggacttccagaagaacaaatgaagtc	4407
Db	4104	CCCCGGTGACAAACCCCAAGAGATGGCTCTCAGGTGAGCTTCCAGAAACCATATGAAGTG	4163
QY	4408	acagagaaactactactccggagatgaataatctctcttaccagaatgatatgtgaagaatttc	4466
Db	4164	ACTGGCGTACACCCAGGGGGGTTCMAAGCTCTTACACGACTGTGACTGAAGAGTTTC	4223
QY	4468	ctcaactccacgaatcaaatgtgcatactatggactctcttttccaagatggcaaaatga	4527
Db	4224	CTGATCAGACAGCAGGAGAGCGGCACTAGTAGACCTGTGTTTCCAAAGCAAGGTG	4283
QY	4528	aaggttttccaaggaaatcaagaactctctcaaacctgtgtgttgaaactctcttagaccacag	4587
Db	4284	AAGGTGTTCAGGGCAACAGAGAGCTTTCACACCGGTGTGTAAACAGCTGTGAGCCCCC	4343



[illegible]

## RESULT 14

```

US-07-864-004B-3
; Sequence 3, Application US/07864004B
; Patent No. 5364771
;
; GENERAL INFORMATION:
;
; APPLICANT: Lollar, John S.
; APPLICANT: Runge, Marschall S.
; TITLE OF INVENTION: Hybrid Human/Porcine Factor VIII
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Kilpatrick & Cody
; STREET: 1100 Peachtree Street
; City: Atlanta
; STATE: Georgia
; COUNTRY: US
; ZIP: 30309
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/864,004B
; FILING DATE: 07 April 1992
; CLASSIFICATION: 435
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: EMU106
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-815-6508
; TELEFAX: 404-815-6555
;
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9009 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapien
; TISSUE TYPE: Liver
;
; FEATURE:
; NAME/KEY: misc.feature (Domain Structure)
; LOCATION: 5001 . . . 7053
; OTHER INFORMATION: /note= "Equivalent to the A3-C1"
; OTHER INFORMATION: domain"
;
; FEATURE:
; NAME/KEY: misc.feature (Domain Structure)
; LOCATION: 1 . . . 2277
; OTHER INFORMATION: /note= "Equivalent to the A1-A2"
; OTHER INFORMATION: domain"
;
US-07-864-004B-3

```

```

Query Match      49.0%; Score 2336.4; DB 1; Length 9009;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2337; Conservative 0; Mismatches 1; Indels 0; Gaps
270 catcaaatagagctccacacgctctcttcgctgcttgcgattcggcttagtc 329
|||||

```

D	150	CATCAATATAGAGCTGTCACACGCTCTTCTTGCTGCGCTTTGCGAATTCGATTAGTGTG	209
Q	330	caccagaagatactacccctgggtgcagltgaaactgctcaatggactatatagcaagtacat	389
D	210	CACCAGAAGATACACTCTGGGTGCAGTGGAACTGCTATGAGACTATATGCAAGTATGCT	269
Q	390	cggtgaagctccctgttgaacgcaagattctctctagagtgccaaatctttccatcca	449
D	270	CGGTGACCTGCTGTGGACCCAAAGATTTCCTCTAGGTGTCCAAAATCTTTTCATTCA	329
Q	450	caaccacgtctgtcaacaanaagactcgtcttgtagaattcaagttcaaccttccaacat	509
D	330	CACCTCAGTGTGTACAAAAGAAAGACTCGTTTGTAGATTACGGTTCCACTTTTCAACAT	389
Q	510	cgctaagccaagggccacccttgaatgggtctgtctagtgctccatcccaagctgaagttta	569
D	390	CGCTTAACCCAAAGGCCACCTCGGATGGGTCTGTAGTCCATACCATCAGGCTGAGGTTTA	449
Q	570	tgatacagtggtattacacttaagaacatgagcttcccaaccgtctagttctatgctgt	629
D	450	TGATACAGTGTCTATACCTTTAAGAACATAGGCTTCCCATCTGTGTAAGTTCATGCTGT	509
Q	630	tgtgtatctccactcgygaagctctctggaggagctgaaatgatatgaacgaacgaagtaag	689
D	510	TGTGTTTCTCTACTCGAAAGACTTCTGAGGAGCTGAAATATGATGATCAGACCGTCAAG	569
Q	690	ggaagaaagaatgataaagtcctcccttgggtggaaagccatataatgtctgcgaagtcct	749
D	570	GGAGAAAGAAGATGATTAAGTCTTCCCTCGTGGAAAGCATACATATGTTGGCAGGTCCT	629
Q	750	gaaagagaatggttccaaatggcctctgaaccactgtgccttaactactatcatcttctca	809
D	630	GAAAGAAATATGTTCAATGTCCTGACCCACCATGTGCTTACCTACTCATATCTTTCTCA	689
Q	810	tgtgagacctgtgtaaaagacttgaatcaagcccataggagccctactagtatgtatgaga	869
D	690	TGTGACCTTGTTAAAGACTTGTAATTCAGGCTCATTTGAGCCCTTACTAGTATGTAGAGA	749
Q	870	agggagcttgcgcaaggaanaagacacagacctgtgcacaatctatactactcttgcgt	929
D	750	AGGAGTCTGTGCCAAGGAAAGACACAGACTTGGCAAAATTTATATCTTTTGTCTGT	809
Q	930	atttgatgaaagggaagaagtgtgcactcagaacaagaactccttgatgcagataggga	989
D	810	ATTTGATGAAGAGGGAAGGAAAGTTGGCACTCAGTAACAAAGAACTCTTATGACGATAGGGA	869
Q	990	tgtctgacatctgcgtcgggcctcgggcctaaatgtgcacacagttcaatggttatgtaaaggtc	104
D	870	TGCTGCATCTGCTGCGGCTTGCCCTTAATATGCAACACAGTAAAGTGTATGTAAACAGGTC	929
Q	1050	tctgcagagctctgatatggaatgcacaggaatacgaatcgaatgtgcatgtatgtaattgg	110
D	930	TCGTGCCAGTCTGTGATGTGATGCCACAGGAATACATGCTATTGGCATGTGATTGGAATGGG	989
Q	1110	caaccaactccctggaagtgcactccaatatctctgaaaggtcacacatctctgttgaagaca	116
D	990	CACCACTCCCTGAAGTGCACATATTTCCGGAAGGTGCACACTTTCTGTGAGGAACCA	104
Q	1170	tcgcgaaggtctccttgaanaatctgcgaanaaattctcctactagtcgcacaaccttggat	122
D	1050	TCCGCAAGGCTCTCTTGGAAATCTCGCCCAATTAATTTCTTACTGCTCAACACTCTTGGAT	110
Q	1230	ggaacctggaagalltctactgtttgtcatatctcttccacacaacatgatatgcaatga	128
D	1110	GGACCTTGGACAGTTTCTACTGTTTGTATATCTCTCCACACAAATGATGGCATVGA	116
Q	1290	agccttatgtcaaaagtgcagcgtgtccagaagaaaccccaactacgaatgaataaataatga	134
D	1170	AGCTTATATCTCAAGATGACAGCTGTCTCAGAGGAACCCCAACTACGATATAAAATATGA	122
Q	1350	agaagcgggaagataatgataatacttacttgttctgaaatgaaatggttgcaagttta	140
D	1230	AGAAGCGGAAGACTATGATGATATATCTTACTGATTTCTGAAATGATGATGTGTCAGGTTTA	128











Db 2310 |||||CATGACCGCCTTACTGAGGTTTCTAGTTGTGACAAGAACACTGGTGATTATTACGAGGA 2369  
Oy 2490 cagttatgaagataatlcagcatactlgtcgaiaaaaaaacaatlgccattgaaccaagaag 2549  
Db 2370 CAGTTATGAAGATATTCTAGCATACTTCTGAGTAAAAACAATGCCATTGAACCAAGANG 2429  
Oy 2550 ctctcccaagaatccaagacacccctagcactagcnaaagaatctaalgccaccaca 2607  
Db 2430 CTCTCCCAAGAAATCAAGACACCCCTAGCAGTAAGCAAAAGCAATTTAATGCCACCACA 2487

Search completed: November 18, 2001, 05:07:44  
job time: 36066 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 18, 2001, 05:14:47 : Search time 919.72 Seconds  
(without alignments)  
3253.111 Million cell updates/sec

Title: US-09-689-430-1-COPY\_150\_4914

Perfect score: 4765

Sequence: 1 cctcttcaagtaaacagta.....ttgggtcgttcttcgcgac 4765

Scoring table:

IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: N.Geneseq\_0601:\*  
2: /SIDSR/gcgdata/geneseq/geneseqn/NA1980.DAT:\*  
3: /SIDSR/gcgdata/geneseq/geneseqn/NA1981.DAT:\*  
4: /SIDSR/gcgdata/geneseq/geneseqn/NA1982.DAT:\*  
5: /SIDSR/gcgdata/geneseq/geneseqn/NA1983.DAT:\*  
6: /SIDSR/gcgdata/geneseq/geneseqn/NA1984.DAT:\*  
7: /SIDSR/gcgdata/geneseq/geneseqn/NA1985.DAT:\*  
8: /SIDSR/gcgdata/geneseq/geneseqn/NA1986.DAT:\*  
9: /SIDSR/gcgdata/geneseq/geneseqn/NA1987.DAT:\*  
10: /SIDSR/gcgdata/geneseq/geneseqn/NA1988.DAT:\*  
11: /SIDSR/gcgdata/geneseq/geneseqn/NA1989.DAT:\*  
12: /SIDSR/gcgdata/geneseq/geneseqn/NA1990.DAT:\*  
13: /SIDSR/gcgdata/geneseq/geneseqn/NA1991.DAT:\*  
14: /SIDSR/gcgdata/geneseq/geneseqn/NA1992.DAT:\*  
15: /SIDSR/gcgdata/geneseq/geneseqn/NA1993.DAT:\*  
16: /SIDSR/gcgdata/geneseq/geneseqn/NA1994.DAT:\*  
17: /SIDSR/gcgdata/geneseq/geneseqn/NA1995.DAT:\*  
18: /SIDSR/gcgdata/geneseq/geneseqn/NA1996.DAT:\*  
19: /SIDSR/gcgdata/geneseq/geneseqn/NA1997.DAT:\*  
20: /SIDSR/gcgdata/geneseq/geneseqn/NA1998.DAT:\*  
21: /SIDSR/gcgdata/geneseq/geneseqn/NA1999.DAT:\*  
22: /SIDSR/gcgdata/geneseq/geneseqn/NA2000.DAT:\*  
23: /SIDSR/gcgdata/geneseq/geneseqn/NA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4413	92.6	4629	20	AAH88293
2	4411.6	92.6	9354	18	AAH73164
3	4355.4	91.4	4629	16	AAH76016
4	4351.2	91.3	4670	19	AAV23339
5	4330.8	90.9	4999	21	AAH00122
6	4325.6	90.8	9164	20	AAH82259
7	4321.6	90.7	4832	19	AAV19581
8	4321.6	90.7	4832	19	AAV15338
9	4319	90.6	12445	21	AAH49232
10	4259.4	89.4	11933	21	AAH00121
11	4173	87.6	4275	10	AAH90654

12	4171.6	87.5	4275	9	AAH80446	Modified factor VI
13	4160	87.5	4272	9	AAH80447	Modified factor VI
14	4161	87.3	4373	20	AAH82258	Beta-domain delete
15	4152.2	87.1	4830	9	AAH81344	Human Factor VIII-
16	4108.8	86.2	4545	9	AAH80444	Modified factor VI
17	3748	78.7	5035	18	AAH69811	Factor VIII-DB695-
18	3708.6	77.8	4616	9	AAH81545	Human Factor VIII-
19	3678	77.2	5094	21	AAH49231	DNA construct H50/
20	3189.8	66.9	4334	19	AAV12142	Homo sapiens facto
21	3189.8	66.9	4334	20	AAH91195	Porcine factor VII
22	3189.8	66.9	4334	22	AAH90541	CDNA encoding porc
23	2605.4	54.7	4451	19	AAH23288	Synthetic human Fa
24	2336.8	49.1	11846	20	AAH82261	Factor VIII protei
25	2336.4	49.0	9009	14	AAH05185	Human Factor VIII
26	2336.4	49.0	9009	18	AAH61548	Factor VIII:C (Arg
27	2336.4	49.0	9009	19	AAH25810	Human factor VIII
28	2336.4	49.0	9009	20	AAH91162	Human factor VIII
29	2336.4	49.0	9009	22	AAH90508	Human factor VIII
30	2334.8	49.0	6300	17	AAH03571	Factor VIII CDNA.
31	2334.8	49.0	8241	9	AAH81439	Factor VIII CDNA.1
32	2334.8	49.0	8241	9	AAH81096	CDNA sequence enco
33	2334.8	49.0	8975	6	AAH50054	Human factor VIII
34	2334.8	49.0	8975	21	AAH238604	Human full-length
35	2334.8	49.0	9029	22	AAH60309	Human factor VIII
36	2334.8	49.0	9068	19	AAH15359	Human factor VIII
37	2334.8	49.0	9080	19	AAH19580	Human factor III e
38	2333.8	49.0	7053	18	AAH51357	Factor VIII:C codi
39	2333.2	49.0	8967	17	AAH31031	Factor VIII full-1
40	2333.2	49.0	8967	22	AAH87526	Human Factor VIII
41	2333.2	49.0	9009	19	AAH18884	Homo sapiens facto
42	2332.2	48.9	7056	15	AAH06615	Sequence of human
43	2323.6	48.8	7440	7	AAH60689	Sequence encoding
44	2313.8	48.6	12022	20	AAH82260	Factor VIII protei
45	2272.4	47.7	7440	6	AAH50375	DNA sequence enco

#### ALIGNMENTS

RESULT	ID	AAH88293	standard; DNA: 4629 BP.
1	AAH88293		
AC	AAH88293		
DT	24-SEP-1999	(first entry)	
DE	Human Factor VIII with B domain deleted cDNA.		
KW	Adenoviral vector; Factor VIII; Factor IX; clotting factor; treatment;		
KW	haemostatic; haemophilia A; haemophilia B; gene therapy; ss.		
OS	Homo sapiens.		
PN	US935935-A.		
PD	10-AUG-1999.		
PF	07-JUN-1995;	95US-0484891.	
PR	07-JUN-1995;	95US-0484891.	
PR	10-JUN-1993;	93US-0074920.	
PR	25-MAR-1994;	94US-0218335.	
PA	(GENE-) GENETIC THERAPY INC.		
PI	Connelly S, Kaleko M, Smith T;		
PI	WPI; 1999-457617/38.		
PT	Adenoviral vectors useful for treating hemophilia		
PS	Example 1; Column 53-58; 90pp; English.		



XX This invention describes novel adenoviral vectors comprising at least  
CC one DNA sequence encoding a clotting factor (Factor IX or Factor VIII).  
CC The vectors of the invention have haemostatic activity. The vectors are  
CC useful for the treatment of hemophilia A or hemophilia B by gene therapy.  
CC This sequence represents human Factor VIII cDNA which has the B domain  
CC deleted.  
XX  
SQ Sequence 4629 BP; 1319 A; 1050 C; 1015 G; 1245 T; 0 other;

Query Match 92.6%; Score 4413; DB 20; Length 4629;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 4419; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 271 atgcaaatagagctctccacgtctctcttctgtgctcttgcagttcgtcttaagtc 330  
DB 1 atgcaaatagagctctccacgtctctcttctgtgctcttgcagttcgtcttaagtc 60  
QY 331 accagaagatactactcctgtgtcagtggaactgtcatgtggaactatgtcaaatgtatctc 390  
DB 61 acaagaagatactactcctgtgtcagtggaactgtcatgtggaactatgtcaaatgtatctc 120  
QY 391 ggtgagctgctgtgtgagcgcgaagattctcctcctgagtgccaaaattcttccatcaac 450  
DB 121 ggtgagctgctgtgtgagcgcgaagattctcctcctgagtgccaaaattcttccatcaac 180  
QY 451 acctcagctgtgtacaagaagactgtgttgaataatcacggtcacctttcaacatc 510  
DB 181 acctcagctgtgtacaagaagactgtgttgaataatcacggtcacctttcaacatc 240  
QY 511 gctaaagcaagccacctgtgagtgtctgctagtcctaccatccagctgaggttat 570  
DB 241 gctaaagcaagccacctgtgagtgtctgctagtcctaccatccagctgaggttat 300  
QY 571 gatacagtgctattacacttaagaacatggtctccactcctgtcagctcattatgtt 630  
DB 301 gatacagtgctattacacttaagaacatggtctccactcctgtcagctcattatgtt 360  
QY 631 ggtgtaccctacttggaagactctctgaggaagctgaataatgaatgaaccagtaaaagg 690  
DB 361 ggtgtaccctacttggaagactctctgaggaagctgaataatgaatgaaccagtaaaagg 420  
QY 691 gagaagaagaatgaataagcttccctgtgtggaagccaataatgtctgcaagctctg 750  
DB 421 gagaagaagaatgaataagcttccctgtgtggaagccaataatgtctgcaagctctg 480  
QY 751 aagaagaatgtctcaatggtcctcagaccactgtgaccttactactatcttcttcat 810  
DB 481 aagaagaatgtctcaatggtcctcagaccactgtgaccttactactatcttcttcat 540  
QY 811 gtgagacctgttaaaagacttgaaatcagagctcattgtgagccctactactatgtatgaa 870  
DB 541 gtgagacctgtgtaaagacttgaaatcagagctcattgtgagccctactactatgtatgaa 600  
QY 871 gggagctctggccaaggaagaacagacctgtcacaatattatactactttgtgtta 930  
DB 601 gggagctctggccaaggaagaacagacctgtcacaatattatactactttgtgtta 660  
QY 931 tttagagaagaagaaattgtgactcagaagaagaagactcctgtatcagaataggat 990  
DB 661 tttagagaagaagaaattgtgactcagaagaagaagactcctgtatcagaataggat 720  
QY 991 gctgcatctctcgggctctggccttaaaatgcacacagtcactgttataaacaagctct 1050  
DB 721 gctgcatctctcgggctctggccttaaaatgcacacagtcactgttataaacaagctct 780  
QY 1051 ctgccaaggtctgattggaatgcacagaaatcagttactatgtgcatgtgattggaatggc 1110  
DB 781 ctgccaaggtctgattggaatgcacagaaatcagttactatgtgcatgtgattggaatggc 840  
QY 1111 accactcctgaatgtgacctcaatatctcctcgaaggtcacacattcttgtgagaacat 1170  
DB 1111 accactcctgaatgtgacctcaatatctcctcgaaggtcacacattcttgtgagaacat 1170

DB 841 accactcctgaatgtgacctcaatatctcctcgaaggtcacacattcttgtgagaacat 900  
QY 1171 cgcacaggtcctctggaaatctgcaccaataacttcttactgtctcaaaccttctatg 1230  
DB 901 cgcacaggtcctctggaaatctgcaccaataacttcttactgtctcaaacacttctatg 960  
QY 1231 gaacttggacagtttctactgttttgttcatactctctcccaacaatgtatgtgaa 1290  
DB 961 gaacttggacagtttctactgttttgttcatactctctcccaacaatgtatgtgaa 1020  
QY 1291 gcttatgtcaaatgtagaagctgtgtcagaaggaacccacagaaatgaaaataatgaa 1350  
DB 1021 gcttatgtcaaatgtagaagctgtgtcagaaggaacccacagaaatgaaaataatgaa 1080  
QY 1351 gaagcggagaagctatgatatgacttactctgattctgaatgagatgtgtcaggttgat 1410  
DB 1081 gaagcggagaagctatgatatgacttactctgattctgaatgagatgtgtcaggttgat 1140  
QY 1411 gatgacaactctcctcttcttaccaaattcgtcagttgccaagaagcatcctaaact 1470  
DB 1141 gatgacaactctcctcttcttaccaaattcgtcagttgccaagaagcatcctaaact 1300  
QY 1471 tgggtacattacattgtcgtcgaagagaagactgggaactatgtctcctatgtctgcc 1530  
DB 1201 tgggtacattacattgtcgtcgaagagaagactgggaactatgtctcctatgtctgcc 1260  
QY 1531 ccgagtgacaagaattataaagaatcaatatttgacaatgtgcccctcagcgattgtag 1590  
DB 1261 ccgagtgacaagaattataaagaatcaatatttgacaatgtgcccctcagcgattgtag 1320  
QY 1591 aagtaaaaaaagttccgaattatgtgcatacacagatgaacacttgaagctgtgaagct 1650  
DB 1321 aagtaaaaaaagttccgaattatgtgcatacacagatgaacacttgaagctgtgaagct 1380  
QY 1651 attcagcattgacaggaactcttggaacttacttattatggggaagttgagacacatg 1710  
DB 1381 attcagcattgacaggaactcttggaacttacttattatggggaagttgagacacatg 1440  
QY 1711 ttgattatatttaagaatcaacgaagcagacacataacatctacccctcaggaatcat 1770  
DB 1441 ttgattatatttaagaatcaacgaagcagacacataacatctacccctcaggaatcat 1500  
QY 1771 gatgtccgtctcttgtatcaagaagatccaaaagttgtaaaactttggaagattt 1830  
DB 1501 gatgtccgtctcttgtatcaagaagatccaaaagttgtaaaactttggaagattt 1560  
QY 1831 ccaattctgcaggaagaatatctcaatatataatgagccagttgagtagaagttggcca 1890  
DB 1561 ccaattctgcaggaagaatatctcaatatataatgagccagttgagtagaagttggcca 1620  
QY 1891 actaatcagaatcctcgtgtcgtgacccgtatattacttagtttcgttaatatgagaga 1950  
DB 1621 actaatcagaatcctcgtgtcgtgacccgtatattacttagtttcgttaatatgagaga 1680  
QY 1951 gatctagcttcaagaatcattgtgcccctcctcctatctgtcacaagaagaatctgtatgaa 2010  
DB 1681 gatctagcttcaagaatcattgtgcccctcctcctatctgtcacaagaagaatctgtatgaa 1740  
QY 2011 agaggaacacagatcaatgtcagacaagaagatgtcaatcctgtttctgtatgtgag 2070  
DB 1741 agaggaacacagatcaatgtcagacaagaagatgtcaatcctgtttctgtatgtgag 1800  
QY 2071 aaccgaagctgttacctacacagagaatataaacaagcttctccccaatccagctgagatg 2130  
DB 1801 aaccgaagctgttacctacacagagaatataaacaagcttctccccaatccagctgagatg 1860  
QY 2131 cagcttggagatccagaagttcccaagctcacaactcatgtcacagatccaatgtctatgtt 2190  
DB 1861 cagcttggagatccagaagttcccaagctcacaactcatgtcacagatccaatgtctatgtt 1920  
QY 2191 tttagtaattgcaagttgtgaagttgtttgtgataaggtgtgcatactgtgattcctcagg 2250  
DB 1921 tttagtaattgcaagttgtgaagttgtttgtgataaggtgtgcatactgtgattcctcagg 1980







Db	4141	ggagtaactactctcaggagtaaaatctctcttaccagcatgtatgtgaaaggattcttc	4200
Oy	4471	actccaggaatgaagatgagccatgtagactctcttttcaagaatgacaaagtaag	4530
Db	4201	actccaggaatcaagaatgagccatgtagactctctcttttcaagaatgacaaagtaag	4260
Oy	4531	gttttcaggagaaatcaagaactctcttcaaccctgtgtgtaactctctagaaccacgtta	4590
Db	4261	gttttcaggagaaatcaagaactctcttcaaccctgtgtgtaactctctagaaccacgtta	4320
Oy	4591	ctgactcgtactaccttgcgaattcaaccaccagaagtgtggtgacacgaattgacctagaatgt	4650
Db	4321	ctgactcgtactaccttgcgaattcaaccaccagaagtgtggtgacacgaattgacctagaatgt	4380
Oy	4651	gaggtctctggtctgagagcacagaactctactgactcgaagcagattc 4699	
Db	4381	gaggtctctggtctgagagcacagaactctactgaaagtgtgacacattgc 4429	

Result	2	
AA73164		
ID	AA73164	standard; CDNA: 9354 BP.
XX		
AC	AA73164;	
XX		
DT	08-APR-1998	(first entry)
XX		
DE	CDNA encoding human B-domain deleted factor VIII.	
XX		
KW	Post-translational regulatory element; PRE; enhancer II; intronless gene;	
KW	surface antigen gene; cytoplasmic accumulation; targeted delivery;	
KW	near consensus splice sequence; blood coagulation factor; factor VIII;	
KW	factor IX; ss.	
XX		
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	2965..7380
FT		/*tag= "a"
FT	misc_feature	5165..5174
FT		/*tag= "b"
FT		/note= "5', near consensus site"
FT	misc_feature	5695..5703
FT		/*tag= "c"
FT		/note= "5', near consensus site"
FT	misc_feature	6320..6328
FT		/*tag= "d"
FT		/note= "5', near consensus site"
FT	misc_feature	6595..5603
FT		/*tag= "e"
FT		/note= "5', near consensus site"
FT	misc_feature	7045..7053
FT		/*tag= "f"
FT		/note= "5', near consensus site"
FT	misc_feature	7143..7152
FT		/*tag= "g"
FT		/note= "5', near consensus site"
FT	misc_feature	3296..3312
FT		/*tag= "h"
FT		/note= "3', near consensus site"
FT	misc_feature	4798..4817
FT		/*tag= "i"
FT		/note= "3', near consensus site"
FT	misc_feature	5023..5045
FT		/*tag= "j"
FT		/note= "3', near consensus site"
FT	misc_feature	5333..5355
FT		/*tag= "k"
FT		/note= "3', near consensus site"
FT	misc_feature	5520..5538
FT		/*tag= "l"
FT		/note= "3', near consensus site"

```

5604..5632
/*tag= m
/note= "j", near consensus site"
5717..5745
/*tag= n
/note= "j", near consensus site"
6239..6258
/*tag= o
/note= "j", near consensus site"
6658..6682
/*tag= p
/note= "j", near consensus site"
7159..7176
/*tag= q
/note= "j", near consensus site"
7196..7209
/*tag= r
/note= "j", near consensus site"
7289..7315
/*tag= s
/note= "j", near consensus site"
7411..7429
/*tag= t
/note= "j", near consensus site"
7611..8197
/*tag= u
/note= "PRE sequence"

```

PN MO9733994-AL.  
 XX  
 PD 18-SEP-1997.  
 XX  
 PF 10-MAR-1997; 97WO-US03561.  
 XX  
 PR 11-MAR-1996; 96US-0683839.  
 XX  
 PA (IMMU-) IMMUNE RESPONSE CORP.  
 XX  
 PI Bidlingmaler S, I11 CR;  
 DR WP1: 1997-470874/43.  
 XX P-PSDB; AAW23414.  
 XX  
 PT Vector for increased expression of intronless genes - comprises  
 PT intronless gene with at least one near consensus splice sequence, a  
 PT promoter and at least one viral cis-acting post-transcriptional  
 PT regulatory element  
 PS  
 PS Example 1; Pages 21-31; 59pp: English.  
 XX  
 CC The present sequence represents human B-domain deleted factor VIII  
 CC cDNA, and a post-translational regulatory element (PRE) of the  
 CC Hepatitis B virus, which is present 3' of the STOP codon for factor VIII.  
 CC PRE sequences have been shown to function in cis to increase the  
 CC steady-state levels of surface gene transcripts by facilitating  
 CC cytoplasmic accumulation of these transcripts. The present sequence  
 CC is part of a novel vector, comprising an intronless gene containing  
 CC 1 or more near consensus splice sequences operably linked to a  
 CC promoter sequence so that the gene is transcribed in a cell.  
 CC Intronless gene transcripts which contain near consensus splice site  
 CC sequences are believed to get tied up in the nucleus of the cell where  
 CC splicing occurs, rather than being transported to the cytoplasm where  
 CC they can be translated into proteins. The PRE sequences are transcribed  
 CC along with the gene, causing export of the gene transcript from the  
 CC nucleus into the cytoplasm of the cell. The vector can be used  
 CC to increase the expression of an intronless gene containing at least one  
 CC near consensus splice sites, preferably cDNA encoding a blood coagulation  
 CC factor, particularly Factor VIII or IX. The complex allows the targeted  
 CC delivery of the vector to a specific cell, e.g. hepatocytes when the  
 CC ligand is an asialoglycoprotein which binds the asialoglycoprotein  
 CC receptor present on their surface.  
 XX  
 XX Sequence 9354 BP; 2506 A; 2239 C; 2161 G; 2448 T; 0 other;



[illegible]

QY	1286	tgyaagcttatgttcaaaagtacagacgcgttcacagaggaaaccccaactacgaatgnaaaata	1345
Dp	3980	tgyaagcttatgttcaaaagtacagacgcgttcacagaggaaaccccaactacgaatgnaaaata	4039
QY	1346	atgaaagaagcggaaagactatgatagatagtatctctgatacttctgaaatgtagatggtgcagat	1405
Dp	4040	atgaaagaagcggaaagactatgatagatagtatctctgatacttctgaaatgtagatggtgcagat	4099
QY	1406	tgtatgatagacaactctctctcttccattaccnaattcgtcagttgcagaagaatcctcta	1465
Dp	4100	tgtatgatagacaactctctctcttccattaccnaattcgtcagttgcagaagaatcctcta	4159
QY	1466	aaactctgtgtacatcttaactatgtctgtcgtgaaagaggaagacgtggagactatctcccttagc	1525
Dp	4160	aaactctgtgtacatcttaactatgtctgtcgtgaaagaggaagacgtggagactatctcccttagc	4219
QY	1526	tcgcccgcgagatgacaagaagtataaaagttaaatatttgacaatctgacctcagcggatg	1585
Dp	4220	tcgcccgcgagatgacaagaagtataaaagttaaatatttgacaatctgacctcagcggatg	4279
QY	1586	gtagaagaagtacaaaaaagctccgattatgtgcataacagatgaaacctttaagactcgtg	1645
Dp	4280	gtagaagaagtacaaaaaagctccgattatgtgcataacagatgaaacctttaagactcgtg	4339
QY	1646	aagctatctcaagatgaatctcaggaactcttgagaccttactttatggggaaagtgttgagaca	1705
Dp	4340	aagctatctcaagatgaatctcaggaactcttgagaccttactttatggggaaagtgttgagaca	4399
QY	1706	cactgttgcattatatacttaaaagaatccaagcagaacgacacatatacttaacctcagcga	1765
Dp	4400	cactgttgcattatatacttaaaagaatccaagcagaacgacacatatacttaacctcagcga	4459
QY	1766	tcactgattgtccgtccctttgtatccaaggagattaccaaaagtgttaaacatttgaag	1825
Dp	4460	tcactgattgtccgtccctttgtatccaaggagattaccaaaagtgttaaacatttgaag	4519
QY	1826	attttccaattctcgcgcggagaaatactcaaatataaattgagcagttgactgtgaaagatg	1885
Dp	4520	attttccaattctcgcgcggagaaatactcaaatataaattgagcagttgactgtgaaagatg	4579
QY	1886	ggccaactaatctcagatccctcggttgcctcgaacccgcctatactcctagtttcgttaatatg	1945
Dp	4580	ggccaactaatctcagatccctcggttgcctcgaacccgcctatactcctagtttcgttaatatg	4639
QY	1946	agagagatctcagctctcagagactatcgtgcctctctctcactcgtccacaagaagatctgag	2005
Dp	4640	agagagatctcagctctcagagactatcgtgcctctctctcactcgtccacaagaagatctgag	4699
QY	2006	atcaaaagaggaacagataaigtccagacaagaaggaaigtcatccctgtttctgtlatttg	2065
Dp	4700	atcaaaagaggaacagataaigtccagacaagaaggaaigtcatccctgtttctgtlatttg	4759
QY	2066	atgggaacccgaagctgtgtactctcagagagaataataaaagctttctcccaatccagtg	2125
Dp	4760	atgggaacccgaagctgtgtactctcagagagaataataaaagctttctcccaatccagtg	4819
QY	2126	gagtgcagctgtgaagatccagagttcccaagcctccacaactcatgtacagcatcaatgct	2185
Dp	4820	gagtgcagctgtgaagatccagagttcccaagcctccacaactcatgtacagcatcaatgct	4879
QY	2186	atgtttttgatagtttgcagttgtcagtttgtttgcatagaggtgagcaatcgtgtacatc	2245
Dp	4880	atgtttttgatagtttgcagttgtcagtttgtttgcatagaggtgagcaatcgtgtacatc	4939
QY	2246	taagcatctggagcacaagaagacttccctctcgtctctctctctcgtctcgtatataactttcaac	2305
Dp	4940	taagcatctggagcacaagaagacttccctctcgtctctctctctcgtctcgtatataactttcaac	4999
QY	2306	acaaaatgtctatgaaagacacatccacctatcccatcttcagagaaaatgctctca	2365
Dp	5000	acaaaatgtctatgaaagacacatccacctatcccatcttcagagaaaatgctctca	5059



QY	2366	tgccga tggaaaaa cccagjgcta tga ttc tgg ggtgc caa cta cga cttc tgg aca	2425
Db	5060	tg tccga tggaaaaa cccagjgcta tga ttc tgg ggtgc caa cta cga cttc tgg aca	5119
QY	2426	ggagca tggccgc cttacta ctaaa gttctctag ttgtg caa gaa ca cgttgat taa cgt	2485
Db	5120	ggagca tggccgc cttacta ctaaa gttctctag ttgtg caa gaa ca cgttgat taa cgt	5179
QY	2486	aggaca gttata tga aga tatlctcag ca tactgtgc tag taa aaaa caa tgc caa tga aca a	2545
Db	5180	aggaca gttata tga aga tatlctcag ca tactgtgc tag taa aaaa caa tgc caa tga aca a	5239
QY	2546	gaagctctccca gaattca aaga ca cctcag cta ggc aaa gca aat tta tgc aacc	2605
Db	5240	gaagctctccca gaattca aaga ca cctcag cta ggc aaa gca aat tta tgc aacc	5299
QY	2606	caacga tctctga aagccca tcaac cgggaaa taa cttcgt taa cttc cta cgtcaga taa g	2665
Db	5300	caacga tctctga aagccca tcaac cgggaaa taa cttcgt taa cttc cta cgtcaga taa g	5359
QY	2666	aggaaa tctga cta tga tga tta cca taa tca cttgaa cttga aatgaa gaatttg aca ttt	2725
Db	5360	aggaaa tctga cta tga tga tta cca taa tca cttgaa cttga aatgaa gaatttg aca ttt	5419
QY	2726	atga tga aga taa tga aat tca ga gccccgc agc cttca aaaa gaaa ca cga ca cta ttt a	2785
Db	5420	atga tga aga taa tga aat tca ga gccccgc agc cttca aaaa gaaa ca cga ca cta ttt a	5479
QY	2786	tgcgtcga gttg aaga gctc tgc gata tga tga tga tga cta cccca ca tgc tta aaga	2845
Db	5480	tgcgtcga gttg aaga gctc tgc gata tga tga tga tga cta cccca ca tgc tta aaga	5539
QY	2846	acaga ggtca ga gttgc gacgtgc cctca gttca aaga gttgtttcc aaga aat tta cgt	2905
Db	5540	acaga ggtca ga gttgc gacgtgc cctca gttca aaga gttgtttcc aaga aat tta cgt	5599
QY	2906	atgcgtccttacta cga cctta taa cgttg aaga aactaa tga aca ctttgg aactcgtg	2965
Db	5600	atgcgtccttacta cga cctta taa cgttg aaga aactaa tga aca ctttgg aactcgtg	5659
QY	2966	ggccata taa gaa gaa gtttga aata taa tca tgc tga aacttca aaga aatga agcct	3025
Db	5660	ggccata taa gaa gaa gtttga aata taa tca tgc tga aacttca aaga aatga agcct	5719
QY	3026	ccgcgtccca ttcctca tctca tctca gactta tttctta tga gga aatca aaga gca aagag	3085
Db	5720	ccgcgtccca ttcctca tctca tctca gactta tttctta tga gga aatca aaga gca aagag	5779
QY	3086	caga aactga aaaaaa cttgtca agc taa tga aaaa caaa cta ctttgg aag tgc	3145
Db	5780	caga aactga aaaaaa cttgtca agc taa tga aaaa caaa cta ctttgg aag tgc	5839
QY	3146	aacatcata tgg caa cccca taa ga tga gtttga c tgc caa gcttgg aactat tctcgtg	3205
Db	5840	aacatcata tgg caa cccca taa ga tga gtttga c tgc caa gcttgg aactat tctcgtg	5899
QY	3206	atgttgc aactg gaaaaa aaga tgtgc a cta cga gcttga tga a cccctcgtgc tgc caa	3265
Db	5900	atgttgc aactg gaaaaa aaga tgtgc a cta cga gcttga tga a cccctcgtgc tgc caa	5959
QY	3266	ctaaca ca cttga a cccctgc tca tvgg aaga ca ag tga cag taca gga aat tgc tgc tttc	3325
Db	5960	ctaaca ca cttga a cccctgc tca tvgg aaga ca ag tga cag taca gga aat tgc tgc tttc	6019
QY	3326	tcaacatc tttga tga gaa caaa agcttgc tca cttca cttga aaaa tga gaa ga aactgc a	3385
Db	6020	tcaacatc tttga tga gaa caaa agcttgc tca cttca cttga aaaa tga gaa ga aactgc a	6079
QY	3386	gggc tccctgc ga aata cca gaa tga aaga tccca ctttga aaga aat tca tgc cttca g	3445
Db	6080	gggc tccctgc ga aata cca gaa tga aaga tccca ctttga aaga aat tca tgc cttca g	6139
QY	3446	caatca aatgc taca taa a tga a taca cta cctgcgttga tga a tgc cta gga taa aaga	3505

Dh	6140	caatcaaaaggctacaataaaygaatagacaactacggtcttgaaayggtccagatcaaaagga	61199
Qy	3506	ttcgatgtgatactbgtcctaagaatggtgcagacatgtaaaacatccattatcttttccagt	3565
Dh	6200	ttcgatgtgatactbgtcctaagaatggtgcagacatgtaaaacatccattatcttccagt	6259
Qy	3566	gacatgtgtcaactgttaccgaaaaaagagatataaaatggcaactgtatactctatac	3625
Dh	6260	gacatgtgtcaactgttaccgaaaaaagagatataaaatggcaactgtatactctatac	6319
Qy	3626	caggtgttttttgagacagtgagaaatgtttacatccaaagctggaaattgggoggtggat	3685
Dh	6320	caggtgttttttgagacagtgagaaatgtttacatccaaagctggaaattgggoggtggat	6379
Qy	3686	gacctatttggcagacatctacatgacgtggggaatggacacattttctgtgtcagcaata	3745
Dh	6380	gacctatttggcagacatctacatgacgtggggaatggacacattttctgtgtcagcaata	6439
Qy	3746	aggtgcagactccctctgggaatggctcttgacacatltagaaattccagatcacgt	3805
Dh	6440	aggtgcagactccctctgggaatggctcttgacacatltagaaatttccagatcacgt	6499
Qy	3806	caggaaacatatgtgacagttgggccccaaagcttgccagacttataattcccgatataca	3865
Dh	6500	caggaaacatatgtgacagttgggccccaaagcttgccagacttataattcccgatataca	6559
Qy	3866	atggccttgagacacaaagagaccccttctctgtgatacaagttgatactgttggacacaatga	3925
Dh	6560	atggccttgagacacaaagagaccccttctctgtgatacaagttgatactgttggacacaatga	6619
Qy	3926	ttatctacggcatcaaaagaccaggtgtcccgctacagaagttctccagccctacatctctc	3985
Dh	6620	ttatctacggcatcaaaagaccaggtgtcccgctacagaagttctccagccctacatctctc	6679
Qy	3986	agtttatcatcatgtatagtcttggatgggaagaagtgtgagacattatccgaggaattcca	4045
Dh	6680	agtttatcatcatgtatagtcttggatgggaagaagtgtgagacattatccgaggaattcca	6739
Qy	4046	cttggaaaccttaatggtctcttcttggcaatgtgatatctgatactgtgataaacacataatt	4105
Dh	6740	cttggaaaccttaatggtctcttcttggcaatgtgatatctgatactgtgataaacacataatt	6799
Qy	4106	ttaaacctccaatathatgtctcgaataacatcggtttgacccaactaatatatagtctcga	4165
Dh	6800	ttaaacctccaatathatgtctcgaataacatcggtttgacccaactaatatatagtctcga	6859
Qy	4166	gcaactcttcgcatgagatgtgatavggctgtgatttaaatgatttgcagcaatgccaattgggaa	4225
Dh	6860	gcaactcttcgcatgagatgtgatavggctgtgatttaaatgatttgcagcaatgccaattgggaa	6919
Qy	4226	ttggagagtaaaagcaatactgaatgagacaaatattacgtcttcacatccattacaataatgt	4285
Dh	6920	ttggagagtaaaagcaatactgaatgagacaaatattacgtcttcacatccattacaataatgt	6979
Qy	4286	ttggcacccttgctcctcttcaaaaagctcgacttcaacctccaagaggaggaatggtctga	4345
Dh	6980	ttggcacccttgctcctcttcaaaaagctcgacttcaacctccaagaggaggaatggtctga	7039
Qy	4346	gacctcaggtgtgaataatccaaaagatgtgctgcgaatgtgacttccaagaagacatgaag	4405
Dh	7040	gacctcaggtgtgaataatccaaaagatgtgctgcgaatgtgacttccaagaagacatgaag	7099
Qy	4406	tcaacaagagtaactactacccaaggggataaaatctctgtcttaccagcatgtatavggaaggt	4465
Dh	7100	tcaacaagagtaactactacccaaggggataaaatctctgtcttaccagcatgtatavggaaggt	7159
Qy	4466	ttcccatctccagcagatcaagaatgagccatcagtgactctcttcttccaagatgtgcaag	4525
Dh	7160	ttcccatctccagcagatcaagaatgagccatcagtgactctcttcttccaagatgtgcaag	7219
Qy	4526	taaaagtttttcagggaataatcaagactccttcaacacgtgtgtgaactctctagaccac	4585



```
Db 7220 taaggattttcaggaaatcaagactcttcacacctgtgtgaactctctagaccac 7279
QY 4586 cgttactgacctgcgtactcttgaaatcaacccccgagtttggtgacacagattgccga 4645
Db 7280 cgttactgacctgcgtactcttgaaatcaacccccgagtttggtgacacagattgccga 7339
QY 4646 ggaagagaggtctggtcgtcgaagcacagagaccctactgctcgaagcgaattc 4699
Db 7340 ggaagagaggtctggtcgtcgaagcacagagaccctactgaggtgagccactgc 7393

RESULT 3
AA076016
ID AA076016 standard; cDNA; 4629 BP.
XX
AC AA076016;
XX
DT 20-JUL-1995 (first entry)
XX
DE B-domain deleted Factor-VIII.
XX
KM Factor-VIII; blood-clotting; hemophilia A; gene therapy;
KM adenovirus; vector; ss.
XX
OS Homo sapiens.
XX
PN M09429471-A.
XX
PD 22-DEC-1994.
XX
PF 13-APR-1994; 94MO-US04075.
XX
PR 10-JUN-1993; 93US-0074920.
PR 25-MAR-1994; 94US-0218335.
XX
PA (GENE-) GENETIC THERAPY INC.
XX
PI Connelly S, Kaleko M, Smith T;
XX
DR WPI; 1995-036495/05.
DR P-PSDB; AAR67709.
XX
PT New adenoviral vectors for treatment of haemophilia - contg. a
PT DNA sequence encoding a clotting factor, partic. Factor VIII or
PT Factor IX
XX
PS Disclosure; Fig. 17A-17C; 116pp; English.
XX
CC Human Factor-VIII cDNA, from which the B domain had been deleted, was
CC used to construct recombinant adenovirus vectors that produced
CC therapeutic levels of the clotting factor when administered to an
CC animal host, potentially providing hemophilia A gene therapy.
XX
SQ Sequence 4629 BP; 1318 A; 1051 C; 1018 G; 1242 T; 0 other;

Query Match 91.48; Score 4355.4; DB 16; Length 4629;
Best Local Similarity 99.08; Pred. No. 0;
Matches 4383; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 271 atggaataagagctccacactgctctctctgtgctcttgagatctgtttagtgc 330
Db 1 atggaataagagctccacactgctctctctgtgctcttgagatctgtttagtgc 60
QY 331 accagaatactactcctggtgtcagtggaactgcatgagactatgcaaaagtcatc 390
Db 61 accagaatactactcctggtgtcagtggaactgcatgagactatgcaaaagtcatc 120
QY 391 ggtgagctgctgtgagcgaagaattccctcctagagtgccaaaatctttccattcaac 450
Db 121 ggtgagctgctgtgagcgaagaattccctcctagagtgccaaaatctttccattcaac 180
QY 451 acccagtcgtgtacaaaagactctgttgttagaatcaagtcacacttccaacatc 510
```

```
Db 181 acccagtcgtgtacaaaagactctgttgttagaatcaagtcacacttccaacatc 240
QY 511 gctaaagcgaagccacctgtgagtgtcgtgaagtcctaccatccaggtgagttat 570
Db 241 gctaaagcgaagccacctgtgagtgtcgtgaagtcctaccatccaggtgagttat 300
QY 571 gatacagtggtcattacacttaagaacatggtccctccctgctcaggtcttcagtgctt 630
Db 301 gatacagtggtcattacacttaagaacatggtccctccctgctcaggtcttcagtgctt 360
QY 631 ggtatccctactgaaagctctcgaaggagctgaatagtatgatacgaagccacgaagg 690
Db 361 ggtatccctactgaaagctctcgaaggagctgaatagtatgatacgaagccacgaagg 420
QY 691 gagaagaagaatgataaagctctccctgtgtggaagccatacatgtctgcaggtcctg 750
Db 421 gagaagaagaatgataaagctctccctgtgtggaagccatacatgtctgcaggtcctg 480
QY 751 aaagagaatgtgccaaatggtccttgaccactgtgaccttactactatcttctcat 810
Db 481 aaagagaatgtgccaaatggtccttgaccactgtgaccttactactatcttctcat 540
QY 811 gtgacctgtgtaaaagacttgaaatcaaggtcctcatgtgagccctactagatgtagaa 870
Db 541 gtgacctgtgtaaaagacttgaaatcaaggtcctcatgtgagccctactactatgtagaa 600
QY 871 gggagctgtgccaaggaaagacacagaccttgccaatttactactcttttgcgtga 930
Db 601 gggagctgtgccaaggaaagacacagaccttgccaatttactactcttttgcgtga 660
QY 931 ttgatgaaggaaagatgtgacctctcgaagaacaaagaactccttgatgcaagataggat 990
Db 661 ttgatgaaggaaagatgtgacctctcgaagaacaaagaactccttgatgcaagataggat 720
QY 991 gctcatctgtctcgagcctcgtgccaataatgacacagtcgaatgtgtatgaacagtgct 1050
Db 721 gctcatctgtctcgagcctcgtgccaataatgacacagtcgaatgtgtatgaacagtgct 780
QY 1051 ctgccaggtctgattgtagtgcacagaagaatcagtcattgtgcatgtgattggaatggc 1110
Db 781 ctgccaggtctgattgtagtgcacagaagaatcagtcattgtgcatgtgattggaatggc 840
QY 1111 accactcctgaagtgacatcaatctcctcgaaggccaacatctctgtgaggaacat 1170
Db 841 accactcctgaagtgacatcaatctcctcgaaggccaacatctctgtgaggaacat 900
QY 1171 cgcagaggtctctggaatctgcgccaataactccttactgctcaaacactctgtatg 1230
Db 901 cgcagaggtctctggaatctgcgccaataactccttactgctcaaacactctgtatg 960
QY 1231 gaccttgagacagtttctactgcttctgtcatatctcttccccaacaatgcatgtgaa 1290
Db 961 gaccttgagacagtttctactgcttctgtcatatctcttccccaacaatgcatgtgaa 1020
QY 1291 gcttatgtcaagtagagcagctgtccagaggaaccccacatagaatgaaataatgaa 1350
Db 1021 gcttatgtcaagtagagcagctgtccagaggaaccccacatagaatgaaataatgaa 1080
QY 1351 gaagcgaagactatgatagatcttactatctcgtgaagtggaatgtgtcaggtttat 1410
Db 1081 gaagcgaagactatgatagatcttactatctcgtgaagtggaatgtgtcaggtttat 1140
QY 1411 gatgacaactctcctcttaccatccaaatgtgctcagttgccaagaagcatcctaaact 1470
Db 1141 gatgacaactctcctcttaccatccaaatgtgctcagttgccaagaagcatcctaaact 1200
QY 1471 tgggtacattatcatgtcgtcgtcgaagagagagacgtggagactatgtcccttgtctgc 1530
Db 1201 tgggtacattatcatgtcgtcgtcgaagagagagacgtggagactatgtcccttgtctgc 1260
QY 1531 cccgatgacagaagttaataaagtcgaatatttgacaatgagccctcagcgagattgtagg 1590
```



Dh 1261 ccgagtcagcagaagtlataaagtlcaatattgaaacaaatgcccctcagcugagtggtagg 1320  
QY 1591 aagracaaaaagtcoccatlctatgcatcacacagatgaacaccttaagaactgtgaagct 1650  
Dh 1321 aagtaaaaaaagtcoccatlctatgcatcacacagatgaacaccttaagaactgtgaagct 1380  
QY 1651 attcagcatgaatcaagaatcttgggaccttactttaatggggaagtggagacacactg 1710  
Dh 1381 attcagcatgaatcaagaatcttgggaccttactttaatggggaagtggagacacactg 1440  
QY 1711 ttgttatatttaagaatcaagaacagcagaccataataacttaccctccacagaaactcct 1770  
Dh 1441 ttgttatatttaagaatcaagaacagcagaccataataacttaccctccacagaaactcct 1500  
QY 1771 gatctccgtcccttgcattctcaagagagattaccacaaagggttaaaaacatttgaagattt 1830  
Dh 1501 gatctccgtcccttgcattctcaagagagattaccacaaagggttaaaaacatttgaagattt 1560  
QY 1831 ccaattctgcagagagaataatccaataataatgagacagtgcgttagaagaatggccca 1890  
Dh 1561 ccaattctgcagagagaataatccaataataatgagacagtgcgttagaagaatggccca 1620  
QY 1891 actaaatcagaatcctcgtgctgacccgctattactagtcttgcgttaataatgagaga 1950  
Dh 1621 actaaatcagaatcctcgtgctgaccccgctattactlctgcgttaataatgagaga 1680  
QY 1951 gatctacgtctcagagacatcatctgcctcctcctcatctgctcaagaagatctgtaagtaa 2010  
Dh 1681 gatctacgtctcagagacatcatctgcctcctcctcatctgctcaagaagatctgtaagtaa 1740  
QY 2011 agaggaacacagataatgctcagacagaagaatgctcatccgttcttcgtatcttgatgag 2070  
Dh 1741 agaggaacacagataatgctcagacagaagaatgctcatccgttcttcgtatcttgatgag 1800  
QY 2071 aacccgaagctggtatccctccacagagaataataaagaacttctccccaatcccgctgagtg 2130  
Dh 1801 aacccgaagctggtatccctccacagagaataataaagaacttctccccaatcccgctgagtg 1860  
QY 2131 cagcttagagatccagagattccaagcctccaacatcatgccaagatcatgtgcattgt 2190  
Dh 1861 cagcttagagatccagagattccaagcctccaacatcatgccaagatcatgtgcattgt 1920  
QY 2191 ttctgatatgttgcagctgctcagcttctgttcgactgagagtgccatctgtaacttcaag 2250  
Dh 1921 ttctgatatgttgcagctgctcagcttctgttcgactgagagtgccatctgtaacttcaag 1980  
QY 2251 attggagacagagactgcttctccttctgtcttctctctgatatcacttcaaaaacaaa 2310  
Dh 1981 attggagacagagactgcttctccttctgtcttctctctgatatcacttcaaaaacaaa 2040  
QY 2311 attgctcatgaagacacactacccatctcccatctcagagagaacagtcttcaatgctg 2370  
Dh 2041 attgctcatgaagacacactacccatctcccatctcagagagaacagtcttcaatgctg 2100  
QY 2371 atggaacaaacccaggtctcactgttcctgggtgccaacactcagactcttcggaacagag 2430  
Dh 2101 atggaacaaacccaggtctcactgttcctgggtgccaacactcagactcttcggaacagag 2160  
QY 2431 atgaaccccttactgaagaggttctcagtgttgacagaagaacactggtgattatcagagac 2490  
Dh 2161 atgaaccccttactgaagaggttctcagtgttgacagaagaacactggtgattatcagagac 2220  
QY 2491 agttatgaagatatcttcaagcacttgcctgagtaaaaaacaaatgccaatggaacaaagac 2550  
Dh 2221 agttatgaagatatcttcaagcacttgcctgagtaaaaaacaaatgccaatggaacaaagac 2280  
QY 2551 ttctccacagaatccaagacacctagcacaagaacaaatgaatgccaacccccacaa 2610  
Dh 2281 ttctccacagaatccaagacacctagcacaagaacaaatgaatgccaacccccacaa 2340  
QY 2611 gtcttgaagaagcactcaagcaggaataactgttactactcttgcctgagtaagaagaa 2670  
Dh 2341 gtcttgaagaagcactcaagcaggaataactgttactactcttgcctgagtaagaagaa 2400

QY 2671 attgacttgatgatcacataatcagltgaaatgaaagaagaattttgacatttatgat 2730  
Dh 2401 attgacttgatgatcacataatcagltgaaatgaaagaagaattttgacatttatgat 2460  
QY 2731 gaggatgaanaatcagagcccccagcagcttccaanaagaacagacactatttatgtct 2790  
Dh 2461 gaggatgaanaatcagagcccccagcagcttccaanaagaacagacactatttatgtct 2520  
QY 2791 ggaagtggagagccttgatgattatgagaatgagtgagctcccaactgtttcaagaagaag 2850  
Dh 2521 ggaagtggagagccttgatgattatgagaatgagtgagtgagctcccaactgtttcaagaagaag 2580  
QY 2851 gtcacagatgagcagtgctccctcaggttccaagaagaatgtgtttccacagaatcttactgagtc 2910  
Dh 2581 gtcacagatgagcagtgctccctcaggttccaagaagaatgtgtttccacagaatcttactgagtc 2640  
QY 2911 tcccttactcagcccttatccogtggagaaactaaatgaaactttgggaccccttggggcca 2970  
Dh 2641 tcccttactcagcccttatccogtggagaaactaaatgaaactttgggaccccttggggcca 2700  
QY 2971 tatataagagcagaagttggaagataataatcagtgttaacttccagaanaatcagccttcgct 3030  
Dh 2701 tatataagagcagaagttggaagataataatcagtgtgttaacttccagaanaatcagccttcgct 2760  
QY 3031 cccctatctccttatcttactagccttattctttaatgagaagaatcagagagcagaagcaga 3090  
Dh 2761 cccctatctccttatcttactagccttattctttaatgagaagaatcagagagcagaagcaga 2820  
QY 3091 ccttaagaaaaaacttgcctaagcctaaatgaacccaanaacttcttggaaagtgcacaat 3150  
Dh 2821 ccttaagaaaaaacttgcctaagcctaaatgaacccaanaacttcttggaaagtgcacaat 2880  
QY 3151 catatgagacccactaaagaatgagatttgactgccaagccttggttcttctcgtatgtt 3210  
Dh 2881 catatgagacccactaaagaatgagatttgactgccaagccttggttcttctcgtatgtt 2940  
QY 3211 gacctggaanaaagatgctcactcagccttgaatgagaccccttctcgtctgccaacactaac 3270  
Dh 2941 gacctggaanaaagatgctcactcagccttgaatgagaccccttctcgtctgccaacactaac 3000  
QY 3271 acactggaacccctgctcactgagagagacaaagtgaacagttacaggaattgtctgcttttcaac 3330  
Dh 3001 acactggaacccctgctcactgagagagacaaagtgaacagttacaggaattgtctgctttcaac 3060  
QY 3331 atcttctgaatgaacccaagaagctggttacttcaactgaaataatggaagaagaactgcagagct 3390  
Dh 3061 atcttctgaatgaacccaagaagctggttacttcaactgaaataatggaagaagaactgcagagct 3120  
QY 3391 cccctgcaatctcagatggaagaatcccaattttaagaagaattatcgcttccatgccaatc 3450  
Dh 3121 cccctgcaatctcagatggaagaatcccaattttaagaagaattatcgcttccatgccaatc 3180  
QY 3451 aatgctcatcaaatgagatacacttaccctgcttgatgaatgctcagatcacaagaattcga 3510  
Dh 3181 aatgctcatcaaatgagatacacttaccctgcttgatgaatgctcagatcacaagaattcga 3240  
QY 3511 tggtaatcgtctcagatgagtcaggcagaatgaanaacatccacttatcttaattcaatgagagac 3570  
Dh 3241 tggtaatcgtctcagatgagtcaggcagaatgaanaacatccacttatcttaattcaatgagagac 3300  
QY 3571 gtgtcactgttacaagaaaaaagagagatataaaatgagcactgtataactctcatcagagt 3630  
Dh 3301 gtgtcactgttacaagaaaaaagagagatataaaatgagcactgtataactctcatcagagt 3360  
QY 3631 gtttttgagacagtgugaaatgttaccatccaagaagctggaatttggcgggtggaatgcctc 3690  
Dh 3361 gtttttgagacagtgugaaatgttaccatccaagaagctggaatttggcgggtggaatgcctc 3420  
QY 3691 attggcagacatctaacatgctgagtgagacacacttttctcgtgtgatacagaataaagtgt 3750  
Dh 3421 attggcagacatctaacatgctgagtgagacacacttttctcgtgtgatacagaataaagtgt 3480



QY 3751 cagatccctgggaatggtctctgacacattagagatttccagcttcaggga 3810  
 |||||||  
 Db 3481 cagatccctgggaatggtctctgacacattagagatttccagcttcaggga 3540  
 |||||||  
 QY 3811 caatatggacagtgggcccacaaagtctgacagacttcatatccgatalcaatctgcc 3870  
 |||||||  
 Db 3541 caatatggacagtgggcccacaaagtctgacagacttcatatccgatalcaatctgcc 3600  
 |||||||  
 QY 3871 tggagcaccagagagcccttttttgatgaagtgatgtgttgaccacaagatattt 3930  
 |||||||  
 Db 3601 tggagcaccagagagcccttttttgatgaagtgatgtgttgaccacaagatattt 3660  
 |||||||  
 QY 3931 caggcatcaagaccagaggtgcgcgtccaaagtctccagctccatctccagttt 3990  
 |||||||  
 Db 3661 caggcatcaagaccagaggtgcgcgtccaaagtctccagctccatctccagttt 3720  
 |||||||  
 QY 3991 atcatcatgtatgtctcttgatgggaagaatggcagacttatctgaggaattccacttggga 4050  
 |||||||  
 Db 3721 atcatcatgtatgtctcttgatgggaagaatggcagacttatctgaggaattccacttggga 3780  
 |||||||  
 QY 4051 accttaatgtctcttttgacaaatgtgatctcatctgggataaaacacataatttttaac 4110  
 |||||||  
 Db 3781 accttaatgtctcttttgacaaatgtgatctcatctgggataaaacacataatttttaac 3840  
 |||||||  
 QY 4111 cctccaatattgtctcgataacacgcgtttgcaaccacatcatatagatctgcagcact 4170  
 |||||||  
 Db 3841 cctccaatattgtctcgataacacgcgtttgcaaccacatcatatagatctgcagcact 3900  
 |||||||  
 QY 4171 ctctcgatagagatgtatgggtctgtgatttaaatagtctgacagcatctgcatgggaatggag 4230  
 |||||||  
 Db 3901 ctctcgatagagatgtatgggtctgtgatttaaatagtctgacagcatctgcatgggaatggag 3960  
 |||||||  
 QY 4231 agtaaaagcaatatacaatgacagatgactgcttacttacttacttccaatatgatttgc 4290  
 |||||||  
 Db 3961 agtaaaagcaatatacaatgacagatgactgcttacttacttacttccaatatgatttgc 4020  
 |||||||  
 QY 4291 acctgtctctctcaaaaagctcgacttcaactccacaaagggaaggaatgactgagact 4350  
 |||||||  
 Db 4021 acctgtctctctcaaaaagctcgacttcaactccacaaagggaaggaatgactgagact 4080  
 |||||||  
 QY 4351 cagggtgaataatccaaaagatgtgctgcaagtgtgacttccagaagaacaatgaaatgcaca 4410  
 |||||||  
 Db 4081 cagggtgaataatccaaaagatgtgctgcaagtgtgacttccagaagaacaatgaaatgcaca 4140  
 |||||||  
 QY 4411 ggggttaactactcagaggaatgaataatctctgtctacacagatgattgtggaagggttcttc 4470  
 |||||||  
 Db 4141 ggggttaactactcagaggaatgaataatctctgtctacacagatgattgtggaagggttcttc 4200  
 |||||||  
 QY 4471 atctccagcagatcaaatgagcagatgactctcttttttcagaatggcaagtcaag 4530  
 |||||||  
 Db 4201 atctccagcagatcaaatgagcagatgactctcttttttcagaatggcaagtcaag 4260  
 |||||||  
 QY 4531 gtttttcagggaatcaagaactctctcaaacctgtgtgaactctctctagaccacagttta 4590  
 |||||||  
 Db 4261 gtttttcagggaatcaagaactctctcaaacctgtgtgaactctctctagaccacagttta 4320  
 |||||||  
 QY 4591 ctgagcctgactacttgatcaatcaacccccagagttgtgtgacacagatttgcctggaagatg 4650  
 |||||||  
 Db 4321 ctgagcctgactacttgatcaatcaacccccagagttgtgtgacacagatttgcctggaagatg 4380  
 |||||||  
 QY 4651 gagggtctggtcgtcgaggaacacagacactctactctgagggtgtgacactgc 4699  
 |||||||  
 Db 4381 gagggtctggtcgtcgaggaacacagacactctactctgagggtgtgacactgc 4429

## RESULT 4

AAV23339 standard; DNA; 4670 BP.

AAV23339;

17-AUG-1998 (first entry)

XX

XX

XX

XX

XX

DE Human Factor-VIII gene lacking central B domain.

XX Factor-VIII: blood clotting; human; synthetic gene; codon usage; ss.

XX Homo sapiens.

OS Synthetic.

XX MO9812207-A1.

XX 26-MAR-1998.

XX 18-SEP-1997: 97MO-US16639.

XX 20-SEP-1996: 96US-0717294.

XX (GENO ) GEN HOSPITAL CORP.

XX Haas J, Seed B;

XX WPI: 1998-217200/19.

XX New synthetic eukaryotic gene(s) - in which non-preferred or less

XX preferred codon(s) are replaced to provide high level expression in

XX mammalian cell(s)

XX Example 3; Fig 12; 92pp; English.

XX This gene codes for a human Factor-VIII protein that lacks the

XX central B domain (amino acids 760-1639) of the native protein. In

XX a novel, claimed synthetic gene (see AAV33288), non-preferred or

XX less preferred codons of the native gene are replaced by codons

XX favored by highly expressed human genes to provide high-level

XX expression in mammalian cells. The synthetic gene was assembled

XX from 29 pairs of oligonucleotides (see AAV23340-97) which served as

XX PCR templates. Synthetic genes of the invention (see also

XX AAV33289-91) are used for production of recombinant proteins in

XX mammalian cells at levels of at least 500% of those obtained using

XX the natural genes. They can also be used in gene therapy.

XX Sequence 4670 BP; 1325 A; 1072 C; 1036 G; 1237 T; 0 other;

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX



QY 748 ctgaagaagatggtccaatgacctgaccacatgtgccttaactacatcttct 807  
|||||  
Db 504 ctgaagaagatggtccaatgacctgaccacatgtgccttaactacatcttct 563  
QY 808 catgtgaacctgtttaaagaactgaatccaagacctatggagccctaactgtatga 867  
|||||  
Db 564 catgtgaacctgtttaaagaactgaatccaagacctatggagccctaactgtatga 623  
QY 868 gaaggagatctggccaaggaaagacacagaccttgcaaaaatttaactacttttgc 927  
|||||  
Db 624 gaaggagatctggccaaggaaagacacagaccttgcaaaaatttaactacttttgc 683  
QY 928 gtaattgataagaggaaagtctggacctcagaaacaaagaactccttgatgcagatag 987  
|||||  
Db 684 gtaattgataagaggaaagtctggacctcagaaacaaagaactccttgatgcagatag 743  
QY 988 gatgtcgcatactgtctcgagccctggcctaaatacgcacagtcatagttataacagg 1047  
|||||  
Db 744 gatgtcgcatactgtctcgagccctggcctaaatacgcacagtcatagttataacagg 803  
QY 1048 tctctccaaagtctgatatgtgatgacaggaatcagttctatggcatgtgattggaatg 1107  
|||||  
Db 804 tctctccaaagtctgatatgtgatgacaggaatcagttctatggcatgtgattggaatg 863  
QY 1108 ggcacacactcctggaagtgcataatactcctcgaaggctcacacatctctgtgaggag 1167  
|||||  
Db 864 ggcacacactcctggaagtgcataatactcctcgaaggctcacacatctctgtgaggag 923  
QY 1168 catcgccagagcgtctccttggaatctcgccataaacttcttactatgctcaaacctctg 1227  
|||||  
Db 924 catcgccagagcgtctccttggaatctcgccataaacttcttactatgctcaaacctctg 983  
QY 1228 atggaaccttggaagttctcactgatttgcatactcctctcccaacatgagtcagctg 1287  
|||||  
Db 984 atggaaccttggaagttctcactgatttgcatactcctctcccaacatgagtcagctg 1043  
QY 1288 gaagcttatgtcaaaagttagacagctgtccagaggaaaccccaactacgaattgaaaaaat 1347  
|||||  
Db 1044 gaagcttatgtcaaaagttagacagctgtccagaggaaaccccaactacgaattgaaaaaat 1103  
QY 1348 gaagaagcggaaagactatgatatgatacttactatcttgaaatgagtgtgtcaagttc 1407  
|||||  
Db 1104 gaagaagcggaaagactatgatatgatacttactatcttgaaatgagtgtgtgtcaagttc 1163  
QY 1408 gatgatgacaacactcctccttcttataccaattcgtccagcttgccaagaagacatccctaa 1467  
|||||  
Db 1164 gatgatgacaacactcctccttcttataccaattcgtccagcttgccaagaagacatccctaa 1223  
QY 1468 acttgggtacatctacattgtctgtcgtgaaggaggagactgaggactatgctccctagctcc 1527  
|||||  
Db 1224 acttgggtacatctacattgtctgtcgtgaaggaggagactgaggactatgctccctagctcc 1283  
QY 1528 gcccccgaatgacgaagttataaagtcaaatattgaaacaaatggccctcagcgaattggt 1587  
|||||  
Db 1284 gcccccgaatgacgaagttataaagtcaaatattgaaacaaatggccctcagcgaattggt 1343  
QY 1588 aggaagtacaaaaaagtcccatattatgtgcatacacagatgaaacctttaaagccgttgaa 1647  
|||||  
Db 1344 aggaagtacaaaaaagtcccatattatgtgcatacacagatgaaacctttaaagccgttgaa 1403  
QY 1648 gctatcagaatgaatcagaagactctggagaccttacttattatggggaagtgtgagaaca 1707  
|||||  
Db 1404 gctatcagaatgaatcagaagactctggagaccttacttattatggggaagtgtgagaaca 1463  
QY 1708 cgtgtatataatttaagaatccaagcagacacataatacatctaccctcaaggatc 1767  
|||||  
Db 1464 cgtgtatataatttaagaatccaagcagacacataatacatctaccctcaaggatc 1523  
QY 1768 acgagatgcctcctctgtatctcaagagattaccaaaagtgtaaaacatttgaagat 1827  
|||||  
Db 1524 acgagatgcctcctctgtatctcaagagattaccaaaagtgtaaaacatttgaagat 1583

QY 1828 ttccaattctgcagaggaaataatccaatataaatggaacagtgcataagaagtgg 1887  
|||||  
Db 1584 ttccaattctgcagaggaaataatccaatataaatggaacagtgcataagaagtgg 1643  
QY 1888 ccaactaatatagatctcctggtgcctgacccgcattacttacttactgttgcgttaataagg 1947  
|||||  
Db 1644 ccaactaatatagatctcctggtgcctgacccgcattacttacttactgttgcgttaataagg 1703  
QY 1948 agagatcagcttcaaggatcattatgacctctctcaactcgtccacaagaagatctgtagat 2007  
|||||  
Db 1704 agagatcagcttcaaggatcattatgacctctctcaactcgtccacaagaagatctgtagat 1763  
QY 2008 caaagaggaaacccaataatgttcagacagaaggaaatgcatccctgtttctgtatttgat 2067  
|||||  
Db 1764 caaagaggaaacccaataatgttcagacagaaggaaatgcatccctgtttctgtatttgat 1823  
QY 2068 gagaacccaagctgttactcctacagaataataaagcttctcccaatccaagcttga 2127  
|||||  
Db 1824 gagaacccaagctgttactcctacagaataataaagcttctcccaatccaagcttga 1883  
QY 2128 gtgcagcttgaggaatccagaagttccaagcctccaacatacgtcacagacataatggctat 2187  
|||||  
Db 1884 gtgcagcttgaggaatccagaagttccaagcctccaacatacgtcacagacataatggctat 1943  
QY 2188 gttttgatagtttgagctgtgcaggttgttgcataaggtgtgcatactgtatctta 2247  
|||||  
Db 1944 gttttgatagtttgagctgtgcaggttgttgcataaggtgtgcatactgtatctta 2003  
QY 2248 agcatitggagacagactgacttcttctgtctctctctctctgatatcacttcaaac 2307  
|||||  
Db 2004 agcatitggagacagactgacttcttctgtctctctctctctgatatcacttcaaac 2063  
QY 2308 aaaaatgtctataagacacacactcaacctatcccatctccaaggagaagactgtctcatg 2367  
|||||  
Db 2064 aaaaatgtctataagacacacactcaacctatcccatctccaaggagaagactgtctcatg 2123  
QY 2368 tgcattgaaanaacccaagttctatgtatctcgggtgtgccacaactcagaacttcggaacga 2427  
|||||  
Db 2124 tgcattgaaanaacccaagttctatgtatctcgggtgtgccacaactcagaacttcggaacga 2183  
QY 2428 ggcattgacccgcttactcgaaggttctcagttgtgacaagaacactgtgtattatccgg 2487  
|||||  
Db 2184 ggcattgacccgcttactcgaaggttctcagttgtgacaagaacactgtgtattatccgg 2243  
QY 2488 gacagttatgagaatatttcaagatacttctgtagttaaataacatgccaattgacaacaga 2547  
|||||  
Db 2244 gacagttatgagaatatttcaagatacttctgtagttaaataacatgccaattgacaacaga 2303  
QY 2548 agcttctcccaagaattccaagacaccttagcactagagcaaaaagcaatttaatgcccacca 2607  
|||||  
Db 2304 agcttctcccaagaattccaagacaccttagcactagagcaaaaagcaatttaatgcccacca 2363  
QY 2608 ccaagcttggaaagcccatccaagggaaataactgtaactactccttcaagttagatccaag 2667  
|||||  
Db 2364 ccaagcttggaaagcccatccaagggaaataactgtaactactccttcaagttagatccaag 2423  
QY 2668 gaaattgactatgtatatacatatcagttgaaatggaaggagaatttggacattat 2727  
|||||  
Db 2424 gaaattgactatgtatatacatatcagttgaaatggaaggagaatttggacattat 2483  
QY 2728 gatagaatgaataatccaagagcccccagcttccaagaagaacagacacttttatt 2787  
|||||  
Db 2484 gatagaatgaataatccaagagcccccagcttccaagaagaacagacacttttatt 2543  
QY 2788 gctgcagttggagagctcctggaattatggagatgagtgccccaatgttctcaagaaac 2847  
|||||  
Db 2544 gctgcagttggagagctcctggaattatggagatgagtgccccaatgttctcaagaaac 2603  
QY 2848 agggctcaaggtgtgcagtgctccctcagttccaagaaagtgttttccaaggaatttactgat 2907  
|||||  
Db 2604 agggctcaaggtgtgcagtgctccctcagttccaagaaagtgttttccaaggaatttactgat 2663  
QY 2908 ggtccttactcagcccttatacggtgaggaactaaatgaacatttggagctcctcgagg 2967



|||||  
Db 2664 ggtccttactcagcccttataccgtgagaaactgaaacttcttgagcctctgagg 2723  
QY 2968 ccatataaagagcagaagtttgaagataataatcatgtgaacttccagaatcagcctct 3027  
Db 2724 ccatataaagagcagaagtttgaagataataatcatgtgaacttccagaatcagcctct 2783  
QY 3028 gctccacttctctctatcttagccttattcttatgaggaagatcagaggcaggagca 3087  
Db 2784 gctccacttctctctatcttagccttattcttatgaggaagatcagaggcaggagca 2843  
QY 3088 gaacctagaanaaaacttctgtaagcctaatagaaaccaaacttacttcttgaaagtcaaa 3147  
Db 2844 gaacctagaanaaaacttctgtaagcctaatagaaaccaaacttacttcttgaaagtcaaa 2993  
QY 3148 catcataatgacacccactaaagaatgtgaactgcaaaagcctggtgcttattctctgat 3207  
Db 2904 catcataatgacacccactaaagaatgtgaactgcaaaagcctggtgcttattctctgat 2963  
QY 3208 gtgacactggaanaagatgtgcaactcagcctgaattgagaccctctggtctgcaact 3267  
Db 2964 gtgacactggaanaagatgtgcaactcagcctgaattgagaccctctggtctgcaact 3023  
QY 3268 aacacactgaaacctgtctcatgaggagacaagtgaacagtaacggaattgctctgttttc 3327  
Db 3024 aacacactgaaacctgtctcatgaggagacaagtgaacagtaacggaattgctctgttttc 3083  
QY 3328 accacttctgataagaccaaagctgtgaactcaactgtaaaatctgaaagaactcagag 3387  
Db 3084 accacttctgataagaccaaagctgtgaactcaactgtaaaatctgaaagaactcagag 3143  
QY 3388 gctcccgcaatatccagatggaagatcccaactttaagaagaattcgtcttcacgca 3447  
Db 3144 gctcccgcaatatccagatggaagatcccaactttaagaagaattcgtcttcacgca 3203  
QY 3448 atcaatggtctacataatgatactacactgtgcttaatggtctcagatcaaaagatt 3507  
Db 3204 atcaatggtctacataatgatactacactgtgcttaatggtctcagatcaaaagatt 3263  
QY 3508 cgaatggtatctgctcacaagcagatggcagaatgaaacatccatctatcttaagtga 3567  
Db 3264 cgaatggtatctgctcacaagcagatggcagaatgaaacatccatctatcttaagtga 3323  
QY 3568 catgtgtcactgtgacgaaaaaagagatataaaatggcactgtcaatcctatcca 3627  
Db 3324 catgtgtcactgtgacgaaaaaagagatataaaatggcactgtcaatcctatcca 3383  
QY 3628 ggtgttttggagacgtgaaatgttaacatccaaagctggaatttggcgggtggaatgc 3687  
Db 3384 ggtgttttggagacgtgaaatgttaacatccaaagctggaatttggcgggtggaatgc 3443  
QY 3688 cttaattggcagcatctacatgtctggaatgagcaacttctctgtgtacagcaataag 3747  
Db 3444 cttaattggcagcatctacatgtctggaatgagcaacttctctgtgtacagcaataag 3503  
QY 3748 tgttaagactcccgaggaaatggtctctggaacacatagagattttagatttaacgcttca 3807  
Db 3504 tgttaagactcccgaggaaatggtctctggaacacatagagattttagatttaacgcttca 3563  
QY 3808 ggaacaataatgagacgttggcccccaaaagcttggccagacttcaatctccggatcaatcaat 3867  
Db 3564 ggaacaataatgagacgttggcccccaaaagcttggccagacttcaatctccggatcaatcaat 3623  
QY 3868 gctctggagcaccaagggcccttctcttgatcaaggttgatctgttggccaatgatac 3927  
Db 3624 gctctggagcaccaagggcccttctcttgatcaaggttgatctgttggccaatgatac 3683  
QY 3928 attcaggcatcaagacccaaggtgtcccgtaagaagtcttccaacccctcatcatctcag 3987  
Db 3684 attcaggcatcaagacccaaggtgtcccgtaagaagtcttccaacccctcatcatctcag 3743  
QY 3988 ttatactcatgtatagttcttgatggaagaagatgagacattctcagagaatctccact 4047  
|||||

Db 3744 ttatactcatgtatagttcttgatggaagaagatgagacattctcagagaatctccact 3803  
QY 4048 ggaaccttaatggtcttctcttggaacatgttgatctcatctgggataaaacacaaattttc 4107  
Db 3804 ggaaccttaatggtcttctcttggaacatgttgatctcatctgggataaaacacaaattttc 3863  
QY 4108 aacctccaattatgtctgcatacatccggttggaaccaactctatlaagatctgcagc 4167  
Db 3864 aacctccaattatgtctgcatacatccggttggaaccaactctatlaagatctgcagc 3923  
QY 4168 actctgcacatgagatgtgagctgtgatttaaatagtttgcaagatgccaattgggaatg 4227  
Db 3924 actctgcacatgagatgtgagctgtgatttaaatagtttgcaagatgccaattgggaatg 3983  
QY 4228 gagaataaagcaatatacatagatgcacagattactgtcttcaatctcttaccataatgttt 4287  
Db 3984 gagaataaagcaatatacatagatgcacagattactgtcttcaatctcttaccataatgttt 4043  
QY 4288 gccacctgtctctcttcaaaaagctcgacttcaactcccaaggagagtaatgcttgaga 4347  
Db 4044 gccacctgtctctcttcaaaaagctcgacttcaactcccaaggagagtaatgcttgaga 4103  
QY 4348 cctcaggctgaataatccaaaagatggctgcaagtggacttccagaagacaaatgaagctc 4407  
Db 4104 cctcaggctgaataatccaaaagatggctgcaagtggacttccagaagacaaatgaagctc 4163  
QY 4408 acagagatgaactactcagggagataaatctctgcttaccagatgtatgtgaagagcttc 4467  
Db 4164 acagagatgaactactcagggagataaatctctgcttaccagatgtatgtgaagagcttc 4223  
QY 4468 ctcatctccagcaatgaagatggtcacaatggtgactctcttcttccagaatgagcaagta 4527  
Db 4224 ctcatctccagcaatgaagatggtcacaatggtgactctcttcttccagaatgagcaagta 4283  
QY 4528 aaggcttctcagggaataatcaagactccttcaacactgtgtgaactcctctagaaccacg 4587  
Db 4284 aaggcttctcagggaataatcaagactccttcaacactgtgtgaactcctctagaaccacg 4343  
QY 4588 ttactgactcgctacactcttcgaattcaccocccaggttgggtgacacagatgtgacctgag 4647  
Db 4344 ttactgactcgctacactcttcgaattcaccocccaggttgggtgacacagatgtgacctgag 4403  
QY 4648 atgagatctctgggtctgcagacagagacctctactgaggttcgagcagatgc 4699  
Db 4404 atgagatctctgggtctgcagacagagacctctactgaggttcgagcagatgc 4455  
  
RESULT 5  
AAD00122  
ID AAD00122 standard; DNA: 4999 BP.  
AC AAD00122;  
DT 31-JUL-2000 (first entry)  
XX  
DE Recombinant adeno associated vector construct. pVmA.1cF8deltaB.  
XX  
KW Recombinant Adeno Associated Vector; rAAV; pVmA.1cF8deltaB; promoter;  
KW human factor VIII; hFVIII; EFlaIpha; human elongation factor-1aIpha;  
KW human growth hormone; hGH; ITR; inverted terminal repeat; haemophilias;  
KW gene therapy; ds.  
XX  
OS Adeno associated virus.  
XX  
PN WO200023116-A1.  
XX  
PD 27-APR-2000.  
XX  
PF 19-OCT-1999; 99NO-US24495.  
XX  
PR 20-OCT-1998; 98US-0104994.  
PR 24-MAR-1999; 99US-0125974.  
PR 30-JUL-1999; 99US-0364862.



```
XX (AVIG-) AVIGEN INC.
PA
XX Couto LB, Colosi PC;
PI
XX WPI: 2000-339536/29.
DR
XX New recombinant adenovirus-associated vector, useful for gene therapy
PT to treat hemophilia, comprises at least a portion of Factor VIII
PT operably linked to control sequence -
XX
XX Example 9, Fig 6, 92pp; English.
XX
XX The present DNA sequence is a recombinant adenovirus-associated vector,
CC (rAAV) construct pYm4.lcf8delab. This expression vector comprises the
CC promoter, first intron (-573 to +985) of human elongation factor-1alpha
CC (EF1alpha) gene, human Factor VIII coding sequence (hVIII) and a
CC polyadenylation signal from human growth hormone (hGH). This sequence is
CC inserted between the AAV inverted terminal repeat (ITR) regions. The
CC hVIII coding region comprises the heavy chain gene segment with the
CC first 57 bp encoding the 19 amino acid signal peptide, A1 and A2 domains
CC and 5 amino acids from the N-terminus of the B domain. The light chain
CC segment comprises the C-terminal 85 amino acids of B domain and the A3,
CC C1 and C2 domains. Both the heavy and light chain segments are cloned
CC into the same plasmid separated by 42 nucleotides coding for 14 residues
CC of the B domain, that is deleted. This plasmid is operably linked to
CC control sequences, that directs the transcription and translation of the
CC Factor VIII gene. The adeno-associated viral vectors are used for gene
CC therapy to treat haemophilia. This method allows prolonged expression of
CC therapeutic levels of Factor VIII in vivo. The rAAV are used for gene
CC therapy, because of their broad host range, safety profile and duration
CC of expression in the infected hosts.
XX
XX Sequence 4999 BP: 1375 A; 1151 C; 1155 G; 1318 T; 0 other;
SQ
Query Match 90.9%; Score 4330.8; DB 21; Length 4999;
Best Local Similarity 99.0%; Pred. NO. 0;
Matches 4384; Conservative 0; Mismatches 2; Indels 42; Gaps 1;
QY 266 ccacatgcaatagagctccaccctgctctctctctgctgctcttgcatctgcttca 325
DB 411 ccacatgcaatagagctccaccctgctctctctctgctgctcttgcatctgcttca 470
QY 326 gtgcacacagaagatctactcgtgggtgagtcgagtcgacatgaggaactatgcaagtg 385
DB 471 gtgcacacagaagatctactcgtgggtgagtcgagtcgacatgaggaactatgcaagtg 530
QY 386 atctcgtgagctgctgtgagcgaagaattccctccctagagtgccaaaattttccat 445
DB 531 atctcgtgagctgctgtgagcgaagaattccctccctagagtgccaaaattttccat 590
QY 446 tcaacacctgctgctgacaaaagaagactcgttctgtagaattccaggttccattca 505
DB 591 tcaacacctgctgctgacaaaagaagactcgttctgtagaattccaggttccattca 650
QY 506 acatcgtctaagccaagcaccctgagtcgtgtagtgcctaccatccagctgagtg 565
DB 651 acatcgtctaagccaagcaccctgagtcgtgtagtgcctaccatccagctgagtg 710
QY 566 ttattgatacgtggtatctacttaacttaagaacatggtctccatccctgtaattcatg 625
DB 711 ttattgatacgtggtatctacttaacttaagaacatggtctccatccctgtaattcatg 770
QY 626 ctgtgtgtatctctactcgaagaagcttctgagggagctggaataatgatacagacagtc 685
DB 771 ctgtgtgtatctctactcgaagaagcttctgagggagctggaataatgatacagacagtc 830
QY 686 aagaaggaaagaaatgataaagttctcctggtggaagccatacatatgctgcaagg 745
DB 831 aagaaggaaagaaatgataaagttctcctggtggaagccatacatatgctgcaagg 890
QY 746 tccgaagaagaatggttccaatgctctgacccactgtgcttaccatacatattctt 805
```

```
DB 891 tccitgaagaagaatggttccaatgctctgacccactgtgcttaccatacatattctt 950
QY 806 ctcatgtgacctgtgtataaagaactggaattcagggcctattggagccctactagtgta 865
DB 951 ctcatgtgacctgtgtataaagaactggaattcagggcctattggagccctactagtgta 1010
QY 866 gagaaggagtcgtgccaaggaagaagacacagacctgtgcacaaatttatactacttttg 925
DB 1011 gagaaggagtcgtgccaaggaagaagacacagacctgtgcacaaatttatactacttttg 1070
QY 926 ctgtatttgatgaaggaaagtttgctactcagaacaaagaactccttgatgagata 985
DB 1071 ctgtatttgatgaaggaaagtttgctactcagaacaaagaactccttgatgagata 1130
QY 986 gggatgctgcactctctcggcctcgtgccaataatgcacagtcgaatggtatgaaca 1045
DB 1131 gggatgctgcactctctcggcctcgtgccaataatgcacagtcgaatggtatgaaca 1190
QY 1046 ggtctctcgaagtcgtatgtagtgcacaggaatcagtcctattgcatgtgtgaa 1105
DB 1191 ggtctctcgaagtcgtatgtagtgcacaggaatcagtcctattgcatgtgtgaa 1250
QY 1106 ttgggacacactcctgaatgtcaatcaatctcctcgaagtcacacatttctgtgaga 1165
DB 1251 ttgggacacactcctgaatgtcaatcaatctcctcgaagtcacacatttctgtgaga 1310
QY 1166 accatcgccaagcgctccttggaatcctgcgaataacttctctacgtccaacactct 1225
DB 1311 accatcgccaagcgctccttggaatcctgcgaataacttctctacgtccaacactct 1370
QY 1226 tgaatgaccttgagacagttctactcgtttgtlcatatctctccacacaatgtagca 1285
DB 1371 tgaatgaccttgagacagttctactcgtttgtlcatatctctccacacaatgtagca 1430
QY 1286 ttggaagcttatgtcaaaatgagacagctgtccaaggaaccccaactcgaatgtaaaata 1345
DB 1431 ttggaagcttatgtcaaaatgagacagctgtccaaggaaccccaactcgaatgtaaaata 1490
QY 1406 ttgatgatgacaactcctctccttcaatccaaatcgctcagttgcgaagaagatcccta 1465
DB 1551 ttgatgatgacaactcctctccttcaatccaaatcgctcagttgcgaagaagatcccta 1610
QY 1466 aaacttggtatcaatacttgctgtgaaggaggagtgagctatgctcccttaagtc 1525
DB 1611 aaacttggtatcaatacttgctgtgaaggaggagtgagctatgctcccttaagtc 1670
QY 1526 tcgccccgatatcagaagaattataaaatccaatttgaaacaatgagccctcagcgatg 1585
DB 1671 tcgccccgatatcagaagaattataaaatccaatttgaaacaatgagccctcagcgatg 1730
QY 1586 gtaggaagtacaaaaaagtcgcgattattgycatacacagatgaaacctttaagaccgtg 1645
DB 1731 gtaggaagtacaaaaaagtcgcgattattgycatacacagatgaaacctttaagaccgtg 1790
QY 1646 aagctatcagcatgaatcagaagaatcttggacotttaactttaagggaagtggagaca 1705
DB 1791 aagctatcagcatgaatcagaagaatcttggacotttaactttaagggaagtggagaca 1850
QY 1706 cactgttatataatataagaatcaagaacagaacacataaacaatcaacctcaaggaa 1765
DB 1851 cactgttatataatataagaatcaagaacagaacacataaacaatcaacctcaaggaa 1910
QY 1766 tcaactggtccgtcctcttgatccaaggagattaccaaaaagtgtaaaaactttgaagg 1825
DB 1911 tcaactggtccgtcctcttgatccaaggagattaccaaaaagtgtaaaaactttgaagg 1970
QY 1826 atttccaattcgcgaagaaataatccaataataatgagacgttactgttgaagatg 1885
```



Db 1971 atttccaattctgcagagaaatattccaaataataatgacagtgtctgtagaagt 2030  
QY 1886 ggcacaataatcagatccctcggtgctgacccgtattactcagtttctgttaattg 1945  
Db 2031 ggcacaataatcagatccctcggtgctgacccgtattactcagtttctgttaattg 2090  
QY 1946 agagaatctagctcagacatcttgccctccctccatctgtctaaagaatctgtag 2005  
Db 2091 agagaatctagctcagacatcttgccctccctccatctgtctaaagaatctgtag 2150  
QY 2006 atcaaaaggaaacacagataatgtcagacaagaaggaaatgtcatctgtttctgtattg 2065  
Db 2151 atcaaaaggaaacacagataatgtcagacaagaaggaaatgtcatctgtttctgtattg 2210  
QY 2066 atgagaacccgaagtgtgtactccaagagaaatatacagctttctcccaatccagctg 2125  
Db 2211 atgagaacccgaagtgtgtactccaagagaaatatacagctttctcccaatccagctg 2270  
QY 2126 gagtgcagcttgagaaatccagagttccaaagcctccaacatcagcacagcatcagtgc 2185  
Db 2271 gagtgcagcttgagaaatccagagttccaaagcctccaacatcagcacagcatcagtgc 2330  
QY 2186 atgttttttagtaagtttgcagttgtcagttgttgcatagagtgagctactgtgtacatc 2245  
Db 2331 atgttttttagtaagtttgcagttgttgcatagagtgagctactgtgtacatc 2390  
QY 2246 taagcatgtgagacacagacgtactctctctgtctctctctcctcggtataactccaac 2305  
Db 2391 taagcatgtgagacacagacgtactctctctgtctctctctcctcggtataactccaac 2450  
QY 2306 acaaaatggtctatgaaagacacatccacatccatccatccacagagaacgtgtctca 2365  
Db 2451 acaaaatggtctatgaaagacacatccacatccatccatccacagagaacgtgtctca 2510  
QY 2366 tctcgaatgtaaaaacccaggtctatgatactctgggtgtgccaaactcagacttcogaca 2425  
Db 2511 tctcgaatgtaaaaacccaggtctatgatactctgggtgtgccaaactcagacttcogaca 2570  
QY 2426 gaagcagacacgcgtctatcaaggttctgtgtgtgacaagaacacgtggtattatcag 2485  
Db 2571 gaagcagacacgcgtctatcaaggttctgtgtgtgacaagaacacgtggtattatcag 2630  
QY 2486 aggaacagttaatgaagatattcagacatctgtcgtgataaacaatgycatgtgaacaa 2545  
Db 2631 aggaacagttaatgaagatattcagacatctgtcgtgataaacaatgycatgtgaacaa 2690  
QY 2546 gaagcttctccagaatccaagacacccctagcactagggcaaaagttaatgtgcaccc 2605  
Db 2691 gaagcttctccagaatccaagacacccctagcactagggcaaaagttaatgtgcaccc 2708  
QY 2606 cacaagctctgaaagcgcacatcaacgggaaataactgttactactctcagtcagatcaag 2665  
Db 2709 cacaagctctgaaagcgcacatcaacgggaaataactgttactactctcagtcagatcaag 2768  
QY 2666 aggaatctgactatgatacatatcagttgaaatgaaagaaagatattgtacatt 2725  
Db 2769 aggaatctgactatgatacatatcagttgaaatgaaagaaagatattgtacatt 2828  
QY 2726 atgataaggaatgaaatcagagcccccagcgtcttcaaaagaaacaacgacatattta 2785  
Db 2829 atgataaggaatgaaatcagagcccccagcgtcttcaaaagaaacaacgacatattta 2888  
QY 2786 ttgctgagatgagaaagcgtctgagatatagtgatactgttaactccccaatgttctaagaa 2845  
Db 2889 ttgctgagatgagaaagcgtctgagatatagtgatactgttaactccccaatgttctaagaa 2948  
QY 2846 acagggctcagagatgagcaggtccctcagttcaagaagaatgttttccagaaatttaag 2905  
Db 2949 acagggctcagagatgagcaggtccctcagttcaagaagaatgttttccagaaatttaag 3008  
QY 2906 atggtctcttactcagccttataccgtgagagaactaaatgaaacatttggagctccgtg 2965  
Db 3009 atggtctcttactcagccttataccgtgagagaactaaatgaaacatttggagctccgtg 3068

QY 2966 ggcacataataagagcagaagttggaagataatatcatgtaactttcagaatccagcct 3025  
Db 3069 ggcacataataagagcagaagttggaagataatatcatgtaactttcagaatccagcct 3128  
QY 3026 ctgcgtccctatcccttctatctcagccttattcttcttgaggagatcagaagcagaag 3085  
Db 3129 ctgcgtccctatcccttctatctcagccttattcttcttgaggagatcagaagcagaag 3188  
QY 3086 cagaacctaagaaaaaacttgtccaagccttaatgaaccaaacttacttttgaagaatgc 3145  
Db 3189 cagaacctaagaaaaaacttgtccaagccttaatgaaccaaacttacttttgaagaatgc 3248  
QY 3146 aacatacataggaacccaacaaagatgagtttactgtcagaagctgggtctatttctctg 3205  
Db 3249 aacatacataggaacccaacaaagatgagtttactgtcagaagctgggtctatttctctg 3308  
QY 3206 atgttgaccttggaanaaagatgtcactcagaagccttgatctgagacccctctgtgccaca 3365  
Db 3309 atgttgaccttggaanaaagatgtcactcagaagccttgatctgagacccctctgtgccaca 3368  
QY 3266 ctaacacactgaaacccctgtctcatagggagacaaagtgcagtaacagaaattgtctgtttc 3325  
Db 3369 ctaacacactgaaacccctgtctcatagggagacaaagtgcagtaacagaaattgtctgtttc 3428  
QY 3326 tcaacacttttgatgagacaaaagctggtacttcaactgaaatattggaagaacatgca 3385  
Db 3429 tcaacacttttgatgagacaaaagctggtacttcaactgaaatattggaagaacatgca 3488  
QY 3386 gggctccctgcaatatccagaatgaaagatccacttcaagagaattatcgcgttccatg 3445  
Db 3489 gggctccctgcaatatccagaatgaaagatccacttcaagagaattatcgcgttccatg 3548  
QY 3446 caatcaatgtgctacataatgataacactacctgtgcttaagtgtcagagatccaagga 3505  
Db 3549 caatcaatgtgctacataatgataacactacctgtgcttaagtgtcagagatccaagga 3608  
QY 3506 ttgcagtgatctcgtctcagaatgagcagacatagaacaacatccatctcatcttctcag 3565  
Db 3609 ttgcagtgatctcgtctcagaatgagcagacatagaacaacatccatctcatcttctcag 3668  
QY 3566 gacatggtctcactgttaacgaaaaaagagagatataaaatgagcactgtacaatctctatc 3625  
Db 3669 gacatggtctcactgttaacgaaaaaagagagatataaaatgagcactgtacaatctctatc 3728  
QY 3626 caggtgttttttgagacagtgagaaatgttaccatccaagctgtgaatttggcgggttggaat 3685  
Db 3729 caggtgttttttgagacagtgagaaatgttaccatccaagctgtgaatttggcgggttggaat 3788  
QY 3686 gcttatttggcagagcatctacgtctgagatgagcacaacttttctgtgttataagacata 3745  
Db 3789 gcttatttggcagagcatctacgtctgagatgagcacaacttttctgtgttataagacata 3848  
QY 3746 agtgcagactcccttgaggatgtctcttgacaacattagaagatttccaattacagctt 3805  
Db 3849 agtgcagactcccttgaggatgtctcttgacaacattagaagatttccaattacagctt 3908  
QY 3806 caggagacaatatggaacagtggtgcccacaagctgtgccaagacttcatctccgatacatca 3865  
Db 3909 caggagacaatatggaacagtggtgcccacaagctgtgccaagacttcatctccgatacatca 3968  
QY 3866 atgccttgagacacaaaggaaccccttctctgtgatacgaagtgtgtgtgtgagccaaatga 3925  
Db 3969 atgccttgagacacaaaggaaccccttctctgtgatacgaagtgtgtgtgtgagccaaatga 4028  
QY 3926 ttattacagcatcaagaacccaaggtgtcccgatcagaagttctccagccttataatctctc 3985  
Db 4029 ttattacagcatcaagaacccaaggtgtcccgatcagaagttctccagccttataatctctc 4088  
QY 3986 agtttatcatcatgtatagttctgtatgaggagaaggtgtgcagacttatacgaagaattcca 4045  
Db 4089 agtttatcatcatgtatagttctgtatgaggagaaggtgtgcagacttatacgaagaattcca 4148



```
OY 4046 ctggaacctaaatggtctctcttggcaatgtgatcatctcgggataaacaacatattc 4105
    |||
DB 4149 ctggaacctaaatggtctctcttggcaatgtgatcatctcgggataaacaacatattc 4208
OY 4106 tttaacctcaattatgtctgatatacgcgtttggaccacacatataatgatcga 4165
    |||
DB 4209 ttaacctcaattatgtctgatatacgcgtttggaccacacatataatgatcga 4268
OY 4166 gcaactctcgaatgagttgaatggctgtgaatttaaatgattgcagcatgcatggaa 4225
    |||
DB 4269 gcaactctcgaatgagttgaatggctgtgaatttaaatgattgcagcatgcatggaa 4328
OY 4226 tggagagtaaaagcaatcagaatgcacagatgaatgcttcaacttaccatattgt 4285
    |||
DB 4329 tggagagtaaaagcaatcagaatgcacagatgaatgcttcaacttaccatattgt 4388
OY 4286 ttgcacactgtctcctcctcaaaagctgactcactcacaaggaggagtaatgcttga 4345
    |||
DB 4389 ttgcacactgtctcctcctcctcaaaagctgactcactcacaaggaggagtaatgcttga 4448
OY 4346 gacctcaggtgaataatccaaaagatggtctgcaagtgaactccagaagaacatgaag 4405
    |||
DB 4449 gacctcaggtgaataatccaaaagatggtctgcaagtgaactccagaagaacatgaag 4508
OY 4406 tcacaggaagtaactactcagaaggataaaatctctgtctaccagcatgtatgtgaagagt 4465
    |||
DB 4509 tcacaggaagtaactactcagaaggataaaatctctgtctaccagcatgtatgtgaagagt 4568
OY 4466 tcctcatctcagcagatcagaatggccatcagttgactccttttttaagaatgccaag 4525
    |||
DB 4569 tcctcatctcagcagatcagaatggccatcagttgactccttttttaagaatgccaag 4628
OY 4526 taaaggttttcagggaaatcaagactcctcacaacctgtgtgaactctcagaaccac 4585
    |||
DB 4629 taaaggttttcagggaaatcaagactcctcacaacctgtgtgaactctcagaaccac 4688
OY 4586 cgttactcgaactgcctacactcgaattcaccacccagagttggtgcacagattgcctga 4645
    |||
DB 4689 cgttactcgaactgcctacactcgaattcaccacccagagttggtgcacagattgcctga 4748
OY 4646 ggaatggaagttctggtcgtgagacagacacacacttactgactcagac 4693
    |||
DB 4749 ggaatggaagttctggtcgtgagacagacacacacttactgactcagac 4796

RESULT 6
AA82259
ID AAX82259 standard; cDNA; 9164 BP.
XX
AC AAX82259;
XX
DT 18-AUG-1999 (first entry)
XX
DE Beta-domain deleted Factor VIII protein encoding gene (uncorrected).
XX
KW Factor VIII protein; gene modification; gene therapy; clinical disorder;
KW splicing pattern; RNA processing; gene regulation; beta-domain; human;
KW ss.
XX
OS Homo sapiens.
XX
PN WO929848-A1.
XX
PD 17-JUN-1999.
XX
PF 25-NOV-1998; 98MO-US25354.
XX
PR 16-JAN-1998; 98US-0071596.
XX
PR 05-DEC-1997; 97US-0067614.
XX
PA (IMMU-) IMMUNE RESPONSE CORP.
XX
PI Bidlingmaier S, Gonzales JEN, Ill CR, Yang CO;
```

```
XX WP1: 1999-385602/32.
DR P-PSDB; AAY21675.
XX
PT Genes and vectors exhibiting increased expression and novel splicing
PT patterns, useful for expression of, e.g. beta-domain deleted factor
PT VIII
XX
PS Example 2; Page 79-89; 123bp; English.
XX
CC The invention describes novel genes and vectors exhibiting increased
CC expression and novel splicing patterns. It provides a gene encoding a
CC Factor VIII protein, that comprises one or more consensus or near
CC consensus splice sites which have been corrected to increase expression.
CC The method, DNA sequences and expression vectors can be used to increase
CC the expression of a gene, especially a Factor VIII gene. Genes containing
CC modified 5' and/or 3' untranslated regions have optimized expression
CC levels and tissue-specific expression. The methods are used for
CC identification and correction of consensus splice sites, addition of
CC introns, optimization of 5' and 3' untranslated regions and increase in
CC cytoplasmic RNA accumulation. Hence the DNAs are useful in gene therapy
CC to treat a clinical disorder, to study RNA processing and/or gene
CC regulation. The present sequence represents an uncorrected version of
CC beta-domain deleted factor VIII protein encoding gene sequence (construct
CC pC1-2). This was used to develop a new coding sequence for beta-domain
CC deleted factor VIII protein by correcting the consensus splice sites.
CC
SQ
Sequence 9164 BP; 2451 A; 2170 C; 2081 G; 2462 T; 0 other;

Query Match 90.8%; Score 4325.6; DB 20; Length 9164;
Best Local Similarity 98.8%; Pred. No. 0;
Matches 4383; Conservative 0; Mismatches 9; Indels 42; Gaps 1;

OY 266 ccacatctcaaatagagatgctccacactgctctctcttctgtgctttggatctgcttta 325
    |||
DB 1001 ccacatctcaaatagagatgctccacactgctctctcttctgtgctttggatctgcttta 1060
OY 326 gtgcacacagaagaatactactcctgtgtgcaagtgaactgtatagggatataatgaagt 385
    |||
DB 1061 gtgcacacagaagaatactactcctgtgtgcaagtgaactgtatagggatataatgaagt 1120
OY 386 atctcgtgtgagctgcctgtgtgagcgaagattcctcctcagaatgycacaaatctttccat 445
    |||
DB 1121 atctcgtgtgagctgcctgtgtgagcgaagattcctcctcagaatgycacaaatctttccat 1180
OY 446 tcaaacactcagtcgtgtacacaaaagactcgtttgtagaattcaaggttcacactttca 505
    |||
DB 1181 tcaaacactcagtcgtgtacacaaaagactcgtttgtagaattcaaggttcacactttca 1240
OY 506 acatcgtctaagccaagccaacctgagatgggtctgtcctaggtcctacacacaggctgag 565
    |||
DB 1241 acatcgtctaagccaagccaacctgagatgggtctgtcctaggtcctacacacaggctgag 1300
OY 566 ttatagatacagtggtattacacttaagaacatggtctccatcctgtgaagttcatg 625
    |||
DB 1301 ttatagatacagtggtattacacttaagaacatggtctccatcctgtgaagttcatg 1360
OY 626 ctgtgtgtatcctcactcgaagaactctctgagggagctggaatagatgacacagcgtc 685
    |||
DB 1361 ctgtgtgtatcctcactcgaagaactctctgagggagctggaatagatgacacagcgtc 1420
OY 686 aaaggagaaagaatgataaagttctcctgtgtggaagccatataatgtctgagag 745
    |||
DB 1421 aaaggagaaagaatgataaagttctcctgtgtggaagccatataatgtctgagag 1480
OY 746 tctcgaagaagaatggttccaatggtccttgacccaatgtgcttaactactatattt 805
    |||
DB 1481 tctcgaagaagaatggttccaatggtccttgacccaatgtgcttaactactatattt 1540
OY 806 ctcatgtgacctgtgtaaaagacttgaaatcagggcctcatggagcctctagtatgta 865
    |||
DB 1541 ctcatgtgacctgtgtaaaagacttgaaatcagggcctcatggagcctctagtatgta 1600
```



QY	866	gagaagggagtcctgycgccaagaaaagacacagaccttgcacaattatbactacttttg	925
Db	1601	gagaaaggagtcctgycgccaagaaaagacacagaccttgcacaattatbactacttttg	1660
QY	926	ctgattttgtaagggaaaagcttgccttcagaaacaaagaaccttgccttgatgaaata	985
Db	1661	ctgattttgtaagggaaaagcttgccttcagaaacaaagaaccttgccttgatgaaata	1720
QY	986	ggagatctgcatctgctcggccctgycgctaanaatgcacacagtcaatggttatlaaca	1045
Db	1721	ggagatctgcatctgctcggccctgycgctaanaatgcacacagtcaatggttatlaaca	1780
QY	1046	ggtctctgcgaaggtctgattgtagtgcgcaagaaatcagttcattgtagtattgaa	1105
Db	1781	ggtctctgcgaaggtctgattgtagtgcgcaagaaatcagttcattgtagtattgaa	1840
QY	1106	tgggcacacactcctcgaaagtgcacatcaatactctcgaaagtacacactttctgtgaga	1165
Db	1841	tgggcacacactcctcgaaagtgcacatcaatactctcgaaagtacacactttctgtgaga	1900
QY	1166	acacatcgccagcgctccttggaatacttcgcacaataccttccttaactgctcacaacct	1225
Db	1901	acacatcgccagcgctccttggaatacttcgcacaataccttccttaactgctcacaacct	1960
QY	1226	tgatgagaccttggaacagttttctactctgtttgtgcatactccttcgacacaatgatgga	1285
Db	1961	tgatgagaccttggaacagttttctactctgtttgtgcatactccttcgacacaatgatgga	2020
QY	1286	tggaagcttatgttcaaaagttagacagctgtgtccagaggaaccccaactagaaatgaaaaata	1345
Db	2021	tggaagcttatgttcaaaagttagacagctgtgtccagaggaaccccaactagaaatgaaaaata	2080
QY	1346	atggaagaaggcggaagctaagatgagatccttaccgattcttgcgaatlgatatgtgtcaggt	1405
Db	2081	atggaagaaggcggaagctaagatgagatccttaccgattcttgcgaatlgatatgtgtcaggt	2140
QY	1406	tgtgatagtgacaactcctcctcttcatcacaatctgcgctcagtttcacaaagaagatccta	1465
Db	2141	tgtgatagtgacaactcctcctcttcatcacaatctgcgctcagtttcacaaagaagatccta	2200
QY	1466	aaacttggtgtacattacattatgctgcctgcgaagagaggaactggtgactatgctcccttaagtc	1525
Db	2201	aaacttggtgtacattacattatgctgcctgcgaagagaggaactggtgactatgctcccttaagtc	2260
QY	1526	tgcgccccggatgcaagaagtataaagaagcaatlttgtacaatgycgctcagcgagattg	1585
Db	2261	tgcgccccggatgcaagaagtataaagaagcaatlttgtacaatgycgctcagcgagattg	2320
QY	1586	gtaggaagatcaaaaaaagcttcgaatttatgycatatcacagatgaaacctttaagaactcgtg	1645
Db	2321	gtaggaagatcaaaaaaagcttcgaatttatgycatatcacagatgaaacctttaagaactcgtg	2380
QY	1646	aagctattcagccttgatcaagaaacttttggaaccttttaactatgtaggaanattgaaaca	1705
Db	2381	aagctattcagccttgatcaagaaacttttggaaccttttaactatgtaggaanattgaaaca	2440
QY	1706	cacgttgatataatatthaagaatcaagaacgaagacacatataacatctcacctcaaggaa	1765
Db	2441	cacgttgatataatatthaagaatcaagaacgaagacacatataacatctcacctcaaggaa	2500
QY	1766	tcaactgattgctcgctcctttgtatatcaagaaagttacccaagaaggttaaacatttgaag	1825
Db	2501	tcaactgattgctcgctcctttgtatatcaagaaagttacccaagaaggttaaacatttgaag	2560
QY	1826	atttccaattctgcgccaaggaataatctcaaatataaatgagacagtgccttgagaagatg	1885
Db	2561	atttccaattctgcgccaaggaataatctcaaatataaatgagacagtgccttgagaagatg	2620
QY	1886	ggccaactaaatacagatctcctgggtgcctgcgaacccgcctatbactctagttctgtaataatg	1945
Db	2621	ggccaactaaatacagatctcctgggtgcctgcgaacccgcctatbactctagttctgtaataatg	2680

QY	1946	agagagatctagcttcagagctcaatctaggccctctctctatctctgctacaaagatcttgag	2005
Db	2681	agaagagatctagcttcagagctcaatctaggccctctctctatctctgctacaaagatcttgag	2740
QY	2006	atcaagaaggaacccaataatgttcagacaagaaggaatctcaacctgtttcttgatttg	2065
Db	2741	atcaagaaggaacccaataatgttcagacaagaaggaatctcaacctgtttcttgatttg	2800
QY	2066	atggaagccggaagctctgttaacctacagagagataataaagctctctcccaatccagctg	2125
Db	2801	atggaagccggaagctctgttaacctacagagagataataaagctctctcccaatccagctg	2860
QY	2126	gagtgcagctctgagatccagagttccaaagctccaaatcatatgcaagcatcaatggtc	2185
Db	2861	gagtgcagctctgagatccagagttccaaagctccaaatcatatgcaagcatcaatggtc	2920
QY	2186	atgttttttgatgattttgcagtttgcagttgtttgttgataggttgagcatctgtacatcc	2245
Db	2921	atgttttttgatgattttgcagtttgcagttgtttgttgataggttgagcatctgtacatcc	2980
QY	2246	taagcatctggagccacaagacagcatctctctgtctctctctctgataatactctcaac	2305
Db	2981	taagcatctggagccacaagacagcatctctctgtctctctctctgataatactctcaac	3040
QY	2306	acaaaatggtctatgtaagacacacacccatctccatctccaagagaacatgctctca	2365
Db	3041	acaaaatggtctatgtaagacacacacccatctccatctccaagagaacatgctctca	3100
QY	2366	tgctgatgagaaaacccaaggtcttatgtattctctgggtgacacaaactcagactttggaa	2425
Db	3101	tgctgatgagaaaacccaaggtcttatgtattctctgggtgacacaaactcagactttggaa	3160
QY	2426	gaggcatgagccgctctctctaaaggtttctcattgtgtgtaacaagaacactgtgtatctag	2485
Db	3161	gaggcatgagccgctctctctaaaggtttctcattgtgtgtaacaagaacactgtgtatctag	3220
QY	2486	aggaacagttatgtaagatattctcagcatactctgtcgtagtaaaaaacaaatgcatgaa	2545
Db	3221	aggaacagttatgtaagatattctcagcatactctgtcgtagtaaaaaacaaatgcatgaa	3280
QY	2546	gaagctctccccaagatctcaagacacccctagcacctagcgcaaaagcaatttaatgcca	2605
Db	3281	gaagctctccccaagatctcaagacacccctagcacctagcgcaaaagcaatttaatgcca	3298
QY	2606	caccagctcttgaagccgcacatcaacgggaaataactcgtactactcttcagtcatagata	2665
Db	3299	caccagctcttgaagccgcacatcaacgggaaataactcgtactactcttcagtcatagata	3358
QY	2666	aggaaatctgactatgatatgataccatcatcagtttgaatgtagaagaagatttgacatt	2725
Db	3359	aggaaatctgactatgatatgataccatcatcagtttgaatgtagaagaagatttgacatt	3418
QY	2726	atgattgagatgtaaaatcagagcccccgaagctttcaaaagaacaacgacataattta	2785
Db	3419	atgattgagatgtaaaatcagagcccccgaagctttcaaaagaacaacgacataattta	3478
QY	2786	ttgttgtcagtggaagagagctctgtgattatgtggatgtgtgagctccccacatgtctca	2845
Db	3479	ttgttgtcagtggaagagagctctgtgattatgtggatgtgtgagctccccacatgtctca	3538
QY	2846	acaagagcttcagagatgtagcagctgtccctcagttccaagaagaattgtttccaagaat	2905
Db	3539	acaagagcttcagagatgtagcagctgtccctcagttccaagaagaattgtttccaagaat	3598
QY	2906	atgagctccttactcagcccttataccgtgtagaactaaatgaacatttggagctctcgg	2965
Db	3599	atgagctccttactcagcccttataccgtgtagaactaaatgaacatttggagctctcgg	3658
QY	2966	ggccatatataaagccagagatgtggaagataataatcatgtgtaacttccagaatccag	3025
Db	3659	ggccatatataaagccagagatgtggaagataataatcatgtgtaacttccagaatccag	3718
QY	3026	ctctgtcccatatctctctatctctcagccttatctctatgtagaagaatcagaagcagag	3085



Db 3719 ctgctccctatctcttctatctctagctcttcttctatctgagaagatcagaggaagag 3778  
 QY 3086 cagaacctagaaaaaactctgtctcaagctctaaatgaaccaaactactcttggaagatgc 3145  
 Db 3779 cagaacctagaaaaaactctgtctcaagctctaaatgaaccaaactactcttggaagatgc 3838  
 QY 3146 aacatcatatgagaccactactaaatgagtttgactctgcaaaagcttggtctattctctg 3205  
 Db 3839 aacatcatatgagaccactactaaatgagtttgactctgcaaaagcttggtctattctctg 3898  
 QY 3206 atgttgaacctggaaaaaagatgtgcactccagctcgatctggaacccctctgctgcaca 3265  
 Db 3899 atgttgaacctggaaaaaagatgtgcactccagctcgatctggaacccctctgctgcaca 3958  
 QY 3266 ctcaacacactggaacccctgctcatgtggaagaagatgacagatcaggaatttgcctgtttt 3325  
 Db 3959 ctcaacacactggaacccctgctcatgtggaagaagatgacagatcaggaatttgcctgtttt 4018  
 QY 3326 tcacacatctcttgatgagaccaaaagctgtgactctcactctgaaatattggaagaactgca 3385  
 Db 4019 tcacacatctcttgatgagaccaaaagctgtgactctcactctgaaatattggaagaactgca 4078  
 QY 3386 gggctccctgcaatataccagaatgcagatcccaactttaaagaatatactgcttccatg 3445  
 Db 4079 gggctccctgcaatataccagaatgcagatcccaactttaaagaatatactgcttccatg 4138  
 QY 3446 caatcaatgctcatatataatgatatactactgctctgtaataatgctcaggaatcaaga 3505  
 Db 4139 caatcaatgctcatatataatgatatactactgctctgtaataatgctcaggaatcaaga 4188  
 QY 3506 ttcgatgtgatactctgctcaggaatgggcagcaatggaacacatccatctctatctcagtg 3565  
 Db 4199 ttcgatgtgatactctgctcaggaatgggcagcaatggaacacatccatctctatctcagtg 4258  
 QY 3566 gaacatgtgtcactctgtacggaaaaaaagagatataaaatggccctgtacaatctctatc 3625  
 Db 4259 gaacatgtgtcactctgtacggaaaaaaagagatataaaatggccctgtacaatctctatc 4318  
 QY 3626 caagtgcttttgagacagatggaatgttacaatccaacagctggaatttgccgggtggaat 3685  
 Db 4319 caagtgcttttgagacagatggaatgttacaatccaacagctggaatttgccgggtggaat 4378  
 QY 3686 gacctatctggcagacatctacatgcttggaatgagacacacttctcgtgtacagcaata 3745  
 Db 4379 gacctatctggcagacatctacatgcttggaatgagacacacttctcgtgtacagcaata 4438  
 QY 3746 agtgcataagctccctctgggaatgtctctgacacatctagagatttccagattacagctt 3805  
 Db 4439 agtgcataagctccctctgggaatgtctctgacacatctagagatttccagattacagctt 4498  
 QY 3806 caggaacataatgacagatgggcccccaagctggccagacatctatctcggatcaatca 3865  
 Db 4499 caggaacataatgacagatgggcccccaagctggccagacatctatctcggatcaatca 4558  
 QY 3866 atgctctggaagaccagaagagcccttctctgacataaagttgagatctgttgccacaatga 3925  
 Db 4559 atgctctggaagaccagaagagcccttctctgacataaagttgagatctgttgccacaatga 4618  
 QY 3926 ttaattcagggaatcaagaacccagaggtgcctgcagaaagttctccagccctcaactctc 3985  
 Db 4619 ttaattcagggaatcaagaacccagaggtgcctgcagaaagttctccagccctcaactctc 4678  
 QY 3986 agttatcatcatgtatagcttgatggaagaagttgagacactatcgaagaatccca 4045  
 Db 4679 agttatcatcatgtatagcttgatggaagaagttgagacactatcgaagaatccca 4738  
 QY 4046 ctggaacctaatggtctctctcttggaacatgttggaatcatctggaataaacaataatct 4105  
 Db 4739 ctggaacctaatggtctctctcttggaacatgttggaatcatctggaataaacaataatct 4798  
 QY 4106 ttaacctcctaattatgtctcgatacatccgtttgacccaactcatatagcatctgca 4165

Db 4799 ttaacctcctaattatgtctcgatacatccgtttgacccaactcatatagcatctgca 4858  
 QY 4166 gcaactctcgatgagttgataatggctgtgatttaaatagttgaagatccatctgga 4225  
 Db 4859 gcaactctcgatgagttgataatggctgtgatttaaatagttgaagatccatctgga 4918  
 QY 4226 tggagagtaaaagcaatatacagatgacagatctactgtctcaactactatcaataatgt 4285  
 Db 4919 tggagagtaaaagcaatatacagatgacagatctactgtctcaactactatcaataatgt 4978  
 QY 4286 ttggcactgtgtctctctcaaaaagcttgacttcaactccctcgaaggagagtaagcttga 4345  
 Db 4979 ttggcactgtgtctctctcaaaaagcttgacttcaactccctcgaaggagagtaagcttga 5038  
 QY 4346 gacctcaggtgaataatccaaaagatggtgcaagtggactcccaagaacatgaag 4405  
 Db 5039 gacctcaggtgaataatccaaaagatggtgcaagtggactcccaagaacatgaag 5098  
 QY 4406 tcacagagtaactactcagggagtaaaatctctgcttaccagcatgtatgtaaggat 4465  
 Db 5099 tcacagagtaactactcagggagtaaaatctctgcttaccagcatgtatgtaaggat 5158  
 QY 4466 tcccatctctcagcagatcaagaatggccatcagtggaactctcttttcgaatggcaag 4525  
 Db 5159 tcccatctctcagcagatcaagaatggccatcagtggaactctcttttcgaatggcaag 5218  
 QY 4526 taaagtttttcagggaaatcaaaagactcctcaaacctgtggtgaactcctcagaccac 4585  
 Db 5219 taaagtttttcagggaaatcaaaagactcctcaaacctgtggtgaactcctcagaccac 5278  
 QY 4586 cgttactgactcgtctactccttgaattcaaccccaagatgtggtgacacagattgcctga 4645  
 Db 5279 cgttactgactcgtctactccttgaattcaaccccaagatgtggtgacacagattgcctga 5338  
 QY 4646 ggaatgaggtgtctgtgctgtgcagagcaagacactactactacgcagagatctc 4699  
 Db 5339 ggaatgaggtgtctgtgctgtgcagagcaagacactactactacgaagggtgccaatgc 5392

RESULT 7  
 AAV19581  
 ID AAV19581 standard; cDNA; 4832 BP.  
 XX  
 AC AAV19581;  
 XX  
 DT 06-AUG-1998 (first entry)  
 XX  
 DE Human factor VIII beta-domain deleted SQN deletion cDNA sequence.  
 XX  
 KW Replication defective; recombinant retrovirus; RRV; therapeutic protein; haemophilia; thrombosis; hypercoagulable disorder; liver disease; human; hepatitis; thalassemia; phenylketonuria; lesch-Myhan syndrome; diabetes; cystic fibrosis; Duchenne's Muscular Dystrophy; hypercholesterolemia; hypopituitarism; adenine deaminase deficiency; HIV infection; anaemia; Guacher's syndrome; high blood pressure; Alzheimer's disease; autoimmune; inflammatory disease; factor VIII; ss.  
 KW  
 XX  
 OS Homo sapiens.  
 XX  
 PN W09800541-A2.  
 XX  
 PD 08-JAN-1998.  
 XX  
 PE 02-JUL-1997; 97WO-US11784.  
 XX  
 PR 04-JUN-1997; 97US-0869309.  
 PR 03-JUL-1996; 96US-0645601.  
 PR 13-AUG-1996; 96US-0696381.  
 PA (CHIR ) CHIRON CORP.  
 XX  
 PI Allen JR, Barber JR, Boder M, Chang SMW, Chong K;  
 De LA VEGA D, Depoloni, Greengard J, Hsu DC, Ibanez CF;



PI	Jolly DJ, Lee R, Mittelstaedt DM, Prussak CE, Respass JG;
XX	
DR	WPI: 1998-086966/08.
DR	P-PSDB: AAM46246.
XX	
PT	New replication defective recombinant retro-viruses - which can be
PT	administered to provide long term systemic expression of therapeutic
PT	protein in blood, useful in, e.g. treating hyper-coagulable
XX	disorders
XX	
PS	Example 28; Pages 210-213; 272pp; English.
XX	
CC	This cDNA encodes the beta-domain deleted SON deletion protein of human
CC	factor VIII. This is used in the construction of recombinant retroviral
CC	vectors expressing human factor VIII. The invention provides the
CC	preparation of replication defective recombinant retrovirus (RVV)
CC	expressing a therapeutic protein. The RVV preparation is resistant to
CC	degradation by human complement and is capable of inducing long term
CC	systemic expression of the therapeutic protein when administered
CC	intravenously to a human. The long term systemic expression results in a
CC	measurable level of the therapeutic protein being produced in the blood
CC	of the human for a period of at least 30 days after the administration of
CC	the RVV vector preparation. RVV's can be used for in vivo delivery of
CC	therapeutic protein to treat, e.g. haemophilia A, haemophilia B,
CC	thrombosis, hypercoagulable disorders, liver diseases such as hepatitis,
CC	disorders such as thalassemia, phenylketonuria, Lesch-Nyhan syndrome,
CC	severe combined immunodeficiency (SCID), cystic fibrosis, Duchenne's
CC	Muscular Dystrophy, inherited emphysema, familial hypercholesterolemia,
CC	diabetes, hypopituitarism, adenine deaminase deficiency, alpha-
CC	antitrypsin deficiency, Gaucher's syndrome, anaemia, infections such as
CC	HIV infection, high blood pressure, Alzheimer's disease, autoimmune or
CC	inflammatory disease or graft versus host disease. RVV's are capable of
CC	surviving inactivation in human serum thereby allowing efficient gene
CC	transfer over prolonged periods of time.
XX	
XX	Sequence 4832 BP; 1381 A; 1089 C; 1055 G; 1307 T; 0 other:
S0	
Query Match	90.7%; Score 4321.6; DB 19; Length 4832;
Best Local Similarity	98.8%; Pred. No. 0;
Matches 4379; Conservative	0; Mismatches 9; Indels 42; Gaps 1;
OY	270 catgcaaatagagctctccaccctgctctctctctgtgctcttggcatttgcatttgatgyc 329
DB	71 catgcaaatagagctctccaccctgctctctctctctgtgctcttggcatttgcatttgatgyc 130
OY	330 caccagaagatatactaccctgggtgcagtggaactgtcagcagggaactatagcaaatgcatct 389
DB	131 caccagaagatatactaccctgggtgcagtggaactgtcagcagggaactatagcacaagtgatct 190
OY	390 cgttagctgtcctgtggcgcgaagattccctccctagatgctgccaatactttccattcaa 449
DB	191 cgttagctgtcctgtggcgcgaagattccctccctagatgctgccaatactttccattcaa 250
OY	450 cacctagtcgtgtacaacaaagactctgtttgtagaattcacggtttcaacctttacaact 509
DB	251 cacctagtcgtgtacaacaaagactctgtttgtagaattcaacgagatcaacctttacaact 310
OY	510 cgtcaagcgaagcgccacctgtgagtgcctgtcgtagtgctctacatccagctgaagttta 569
DB	311 cgtcaagcgaagcgccacctgtgagtgcctgtcgtagtgctctacatccagctgaagttta 370
OY	570 tgatacagtggtcatattacacttaagaacaatggtctccatccctgtcagttcttcagtcgt 629
DB	371 tgatacagtggtcatattacacttaagaacaatggtctccatccctgtcagttcttcagtcgt 430
OY	630 tgggtatccctacttgaaagctcttgagggagctgaatatgatgatgacagacagtcct 689
DB	431 tgggtatccctacttgaaagctcttgagggagctgaatatgatgatgacagacagtcct 490
OY	690 gggagaagaagatagataaagcttccctcgtgtggaaagccatacatatgtctggcaggtcct 749
DB	491 gggagaagaagatagataaagcttccctcgtgtggaaagccatacatatgtctggcaggtcct 550

QY	750	gaagagaaatggtccaaatggtcctctgacccactgtgctcttacctactaatatcttctctca	809
Db	551	gaagagaaatggtccaaatggtcctctgacccactgtgctcttacctactaatatcttctctca	610
QY	810	tgtagacctgttaaaaacttgaaattcagaagcccatctggagccctcactagatgtagagga	869
Db	611	tgtagacctgttcaaaaacttgaaattcagaagcccatctggagccctcactagatgtagagga	670
QY	870	aggagagctctgcacaggaagaaagacacagacactctgcacaatttactactctttctgct	929
Db	671	aggagagctctgcacaggaagaaagacacagacactctgcacaatttactactctttctgct	730
QY	930	atttgatgaaagggaaaagtgtggacatccagaacaaaagaactcctctgtatgcagatagga	989
Db	731	atttgatgaaagggaaaagtgtggacatccagaacaaaagaactcctctgtatgcagatagga	790
QY	990	tgctgcacatctctcgtggccctgcgcctaaatgacacagatcgaatgtgttatgttaaacagtc	1049
Db	791	tgctgcacatctctcgtggccctgcgcctaaatgacacagatcgaatgtgttatgttaaacagtc	850
QY	1050	ctctgcacaggtctgattggaatgacacaggaataatcagatctatctggcatgtgatgtagatgg	1109
Db	851	ctctgcacaggtctgattggaatgacacaggaataatcagatctatctggcatgtgatgtagatgg	910
QY	1110	caccacatctctgaaatgtgacatccataatctctcgaaggtacacacatcttctgtgtaggaacca	1169
Db	911	caccacatctctgaaatgtgacatccataatctctcgaaggtacacacatcttctgtgtaggaacca	970
QY	1170	tcgcacagagcgctccctctggaataatctcgcacaaataattcccttactctgtccaaacactcttat	1229
Db	971	tcgcacagagcgctccctctggaataatctcgcacaaataattcccttactctgtccaaacactcttat	1030
QY	1230	ggacctctggaacagttcttactgttctgtgcatatctcttccacacaatgaaatgagatgga	1289
Db	1031	ggacctctggaacagttcttactgttctgtgcatatctcttccacacaatgaaatgagatgga	1090
QY	1290	agcttaatgtaaaatgaaacagcgtgtccagaaggaaccccaactaagaatggaataaatatga	1349
Db	1091	agcttaatgtaaaatgaaacagcgtgtccagaaggaaccccaactaagaatggaataaatatga	1150
QY	1350	agaagcggaagaacatgatatgatctctctgtattctgaaatgggaatgtgtgtccaagtttga	1409
Db	1151	agaagcggaagaacatgatatgatctctctgtattctgaaatgggaatgtgtgtccaagtttga	1210
QY	1410	tgatgacaactctcctctccttattccaatctcgctcagttgcgaagaagcatcctctaaac	1469
Db	1211	tgatgacaactctcctctccttattccaatctcgctcagttgcgaagaagcatcctctaaac	1270
QY	1470	tttgaggtacatattacgtctctctgtaaaaggaagatcgtggactatggtctccttagtccctgc	1529
Db	1271	tttgaggtacatattacgtctctctgtaaaaggaagatcgtggactatggtctccttagtccctgc	1330
QY	1530	cccgcgatgacagaaagtattataaagaatcaatatttgaaacaatggtccctcagcgatgtgag	1589
Db	1331	cccgcgatgacagaaagtattataaagaatcaatatttgaaacaatggtccctcagcgatgtgag	1390
QY	1590	gaagatcacaaaaaagtcgcgatttatgtgcatacacagaatgaaacctttaagactcgtgaagc	1649
Db	1391	gaagatcacaaaaaagtcgcgatttatgtgcatacacagaatgaaacctttaagactcgtgaagc	1450
QY	1650	tattcacacatgaaatccagaatctcttgagacttactatttatggggaagtgtggaagcaact	1709
Db	1451	tattcacacatgaaatccagaatctcttgagacttactatttatggggaagtgtggaagcaact	1510
QY	1710	gttgattatataatttaagaatacaagcaagcgacacataatacaatcctccacgggaataac	1769
Db	1511	gttgattatataatttaagaatacaagcaagcgacacataatacaatcctccacgggaataac	1570
QY	1770	tgatgtccgtctcttgtatctcaaggaagatcaccaaaaaggtgtaaaaacattgaaagattt	1829
Db	1571	tgatgtccgtctcttgtatctcaaggaagatcaccaaaaaggtgtaaaaacattgaaagattt	1630



OY	1830	tcaatcttcgcccggggaatatctcaatatataaavgacagtgactgtagaagcttgggcc	1885
Db	1631	tcacaattccgcccggggaatatctcaatatataaavgacagtgactgtagaagcttgggcc	1690
OY	1890	aactaatcagatcccttcggtgcctgcgaccgctattactctagcttccgttaatavgaaag	1949
Db	1691	aactaaatacagatcccttcggtgcctgcgaccgctattactctagcttccgttaatavgaaag	1750
OY	1950	agatctagcttcagagctccatctgcccctccctccatccctgctacaaagaatctgtagatca	2009
Db	1751	agatctagcttcagagctccatctgcccctccctccatccctgctacaaagaatctgtagatca	1810
OY	2010	aagagggaaacacagataatgtcagacaagaggaatgtcatccgtttctctgtattgtatga	2056
Db	1811	aagagggaaacacagataatgtcagacaagaggaatgtcatccgtttctctgtattgtatga	1870
OY	2070	gaacggaagctgtgtactccaaagaaatatacaagctttctccccaatccagcttgagct	2129
Db	1871	gaacggaagctgtgtactccaaagaaatatacaagctttctccccaatccagcttgagct	1930
OY	2130	gaagctctggagatccagagattcccaagctcccaaatcatgtcacagacatcaatgtgctatgt	2159
Db	1931	gaagctctggagatccagagattcccaagctcccaaatcatgtcacagacatcaatgtgctatgt	1990
OY	2190	ttttgtagatttgacagttgttcagtttgattgttcgaltgaggtgagacatctgtatcctaaag	2249
Db	1991	ttttgtagatttgacagttgttcagtttgattgttcgaltgaggtgagacatctgtatcctaaag	2050
OY	2250	catttgagagacaaagacagactcccttccttcgtctctctccctgatatccctcaaacaa	2309
Db	2051	catttgagagacaaagacagactcccttccttcgtctctctccctgatatccctcaaacaa	2110
OY	2310	aatgtgctatgaaagacacacatccacccctatcccatctccagagagaacatgtctctcatgtc	2369
Db	2111	aatgtgctatgaaagacacacatccacccctatcccatctccagagagaacatgtctctcatgtc	2170
OY	2370	gattggaacaaacccagagctatgtatctctgggtgcccacactcagacttctcggaacagag	2429
Db	2171	gattggaacaaacccagagctatgtatctctgggtgcccacactcagacttctcggaacagag	2230
OY	2430	catgacccgctctatcaaggtttctctagtttgtgcacaagaacatgtgtatatctgcggga	2489
Db	2231	catgacccgctctatcaaggtttctctagtttgtgcacaagaacatgtgtatatctgcggga	2290
OY	2490	cagttatgaaagatatctcaagcatctgtcgtgagtaaaacaatccatctgacccaagaag	2549
Db	2291	cagttatgaaagatatctcaagcatctgtcgtgagtaaaacaatccatctgacccaagaag	2350
OY	2550	ctctccccaagatccaagacacccctagcactaagcaaaagcaalttaatgcccaccac	2609
Db	2351	ctctccccaagatccaagacacccctagcactaagcaaaagcaalttaatgcccaccac	2368
OY	2610	agttctgaaacgcacatcaagggaaataactcgttactacactcttcagctcagatcaagggga	2659
Db	2369	agttctgaaacgcacatcaagggaaataactcgttactacactcttcagctcagatcaagggga	2428
OY	2670	aattgacatgagatataccatcaatcagattgtgaatgaaagaaagatttgcattcttga	2729
Db	2429	aattgacatgagatataccatcaatcagattgtgaatgaaagaaagatttgcattcttga	2488
OY	2730	tgaaggaatgaanaalcagagccccccgcagcttccaagaagaaacacagacataatttatg	2789
Db	2489	tgaaggaatgaanaalcagagccccccgcagcttccaagaagaaacacagacataatttatg	2548
OY	2790	tgcagttgagagaagctcttggtgattatgggattgagtagctcccccacatgttctaagaacag	2849
Db	2549	tgcagttgagagaagctcttggtgattatgggattgagtagctcccccacatgttctaagaacag	2608
OY	2850	ggctcaggaatggagagtgccctccatgaattccaagaagaattgtttccaaggaatttactgttg	2909
Db	2609	ggctcaggaatggagagtgccctccatgaattccaagaagaattgtttccaaggaatttactgttg	2668
OY	2910	cctccttactcaagcccttataccgttgagaagaactaaatgaaacatttggaacttggaactccgtggcc	2969

Db	2669	cccccttaccagcccttataccgctggagaacccaagacacttgggacccctggggcc	2728
Qy	2970	atataaagagcagaagcttggaga taagacatcgtgtaaacttccagaacacagccctcgg	3029
Db	2729	atataaagagcagaagcttggaga taatcactcgtgtaaacttccagaacacagccctcgg	2788
Qy	3030	tcctcatctcttatacttctaagccttattcttataggagaatcagagccaagcgaca	3083
Db	2789	tcctcatctcttatacttctaagccttattcttataggagaatcagagccaagcgaca	2848
Qy	3090	acctgaaaaaaacttctgtaagccttaatgtaacccaacttaacttcttggaaagctgaca	3149
Db	2849	acctgaaaaaaacttctgtaagccttaatgtaacccaacttaacttcttggaaagctgaca	2908
Qy	3150	lcaatactgaccaccaataagaatgagttctgacatgccaagcctgggcttattcttcagat	3209
Db	2909	lcaatactgaccaccaataagaatgagttctgacatgccaagcctgggcttattcttcagat	2968
Qy	3210	tgacctggaaaaaagatctgcatcctaagcctgatactggaccctctctgtctgcacactaa	3269
Db	2969	tgacctggaaaaaagatctgcatcctaagcctgatactggaccctctctgtctgcacactaa	3028
Qy	3270	caacctggaacccctctcaatctgggagaaacaagtgcacgtacaggaattctgctcttctaac	3329
Db	3029	caacctggaacccctctcaatctgggagaaacaagtgcacgtacaggaattctgctcttctaac	3088
Qy	3330	catctcttgaagagccaaagaactgtgtaacttccactgtgaataatagaaagaactgcagggc	3389
Db	3089	catctcttgaagagccaaagaactgtgtaacttccactgtgaataatagaaagaactgcagggc	3148
Qy	3390	tcacctgacaataccaagatggaagaatcccaactttaaagagaatatactgccttccatgacaat	3449
Db	3149	tcacctgacaataccaagatggaagaatcccaactttaaagagaatatactgccttccatgacaat	3208
Qy	3450	caatggtcctaactaatgtgataatacactctggtctatgtaatggtctcagatccaagaattcg	3509
Db	3209	caatggtcctaactaatgtgataatacactctggtctatgtaatggtctcagatccaagaattcg	3268
Qy	3510	atggaatctgtccagatctggcagcagcaatgaaaaacatccacttcatctattcattgaaggaca	3569
Db	3269	atggaatctgtccagatctggcagcagcaatgaaaaacatccacttcatctattcattgaaggaca	3328
Qy	3570	tggtctcactctgaagaaaaaagaggagataaaatctgacactgtacaactctataccagg	3629
Db	3329	tggtctcactctgaagaaaaaagaggagataaaatctgacactgtacaactctataccagg	3388
Qy	3630	tgctttctgaagacagctggaataatgtataccaatccaaagcttggatttggcggcttggaatgcct	3689
Db	3389	tgctttctgaagacagctggaataatgtataccaatccaaagcttggatttggcggcttggaatgcct	3448
Qy	3690	tatttggcgaagacatctaatatctgtagaagacaacatttctctggtctgaagaataaagt	3749
Db	3449	tatttggcgaagacatctaatatctgtagaagacaacatttctctggtctgaagaataaagt	3508
Qy	3750	tcagactcccccctgggaatctgctctctggaacatatagaatttccagatatacagcttcagg	3809
Db	3509	tcagactcccccctgggaatctgctctctggaacatatagaatttccagatatacagcttcagg	3568
Qy	3810	acaatactgacagctgggcccacaaagctgtgcagacttcatatctccggataccaatgc	3869
Db	3569	acaatactgacagctgggcccacaaagctgtgcagacttcatatctccggataccaatgc	3628
Qy	3870	cttggagccaacgaagaccccttctcttgatccaagctggatctggttggccaatgatatt	3929
Db	3629	cttggagccaacgaagaccccttctcttgatccaagctggatctggttggccaatgatatt	3688
Qy	3930	tcacggacataaagcccaaggctggccgctcagaagttctccagccttaactctccagtt	3989
Db	3689	tcacggacataaagcccaaggctggccgctcagaagttctccagccttaactctccagtt	3748
Qy	3990	tatcatcatgtatagctcttatactggagaagctgtgcagactatactgaggaattccacgtg	4049



Dh 3749 tatcatcatgatagatctcttgatcgggaagaagtgagcaactatcaggaataatccactgg 3808  
Qy 4050 aaccttaatggtctctcttggaacatgagatcactcggggaataaacacaatatcttaa 4109  
Dh 3809 aaccttaatggtctctcttggaacatgagatcactcggggaataaacacaatatcttaa 3868  
Qy 4110 cccctcaatlatctgcatcatalcatccgttttgacccaactaatatagatctcgagcac 4169  
Dh 3869 cccctcaatlatctgcatcatalcatccgttttgacccaactaatatagatctcgagcac 3928  
Qy 4170 tcttcgcatgagatctgagctgctgatttaataagtgcagcatgccatgggaatgga 4229  
Dh 3929 tcttcgcatgagatctgagctgctgatttaataagtgcagcatgccatgggaatgga 3988  
Qy 4230 gagtaagcaatcatcagatgacagatctactgctcacttcacttcacaaatgatttgc 4289  
Dh 3989 gagtaagcaatcatcagatgacagatctactgctcacttcacttcacaaatgatttgc 4048  
Qy 4290 caactggtctcctcctcaaaagctcgaactcactcactcaaggaggaatgagagacc 4349  
Dh 4049 caactggtctcctcctcaaaagctcgaactcactcactcaaggaggaatgagagacc 4108  
Qy 4350 tcaagtgataatccaaaagagtgctgcaagtgagctccagagaacaatgaaatgcac 4409  
Dh 4109 tcaagtgataatccaaaagagtgctgcaagtgagctccagagaacaatgaaatgcac 4168  
Qy 4410 agggagtaactactcagggaggaataatctctgctacacagatgatttgaaagggtccct 4469  
Dh 4169 agggagtaactactcagggaggaataatctctgctacacagatgatttgaaagggtccct 4228  
Qy 4470 catctcagagcagtcagaatgagcactcagtgagctctcttttcagaaatgcaagaataa 4529  
Dh 4229 catctcagagcagtcagaatgagcactcagtgagctctcttttcagaaatgcaagaataa 4288  
Qy 4530 ggttttcagagggaatacaagaactcctcacaacctgctgagtaactctcagaaccagctt 4589  
Dh 4289 ggttttcagagggaatacaagaactcctcacaacctgctgagtaactctcagaaccagctt 4348  
Qy 4590 actgactcgcgaactcgaattcaacccacagagtggtgacacagaattccctcgagagat 4649  
Dh 4349 actgactcgcgaactcgaattcaacccacagagtggtgacacagaattccctcgagagat 4408  
Qy 4650 ggaagttctggctgagcagagcacagacactctactgagcagagatc 4699  
Dh 4409 ggaagttctggctgagcagagcacagacactctactgagcagagatc 4458

RESULT 8  
AAV15338 standard; DNA; 4832 BP.  
ID AAV15338;  
XX  
AC AAV15338;  
XX  
DT 20-JUL-1998 (first entry)  
XX  
DE Human Factor VIII SQN deletion mutant DNA.  
XX  
KW Factor VIII; blood clotting; haemophilia A; gene therapy;  
KW retrovirus; vector; human; ss.  
XX  
OS Homo sapiens.  
OS Synthetic.  
XX  
XX Key Location/Qualifiers  
FT CDS 72..445 /\*tag= a  
FT  
XX  
XX W09800542-A2.  
XX  
XX 08-JAN-1998.  
XX  
XX 02-JUL-1997; 97W0-US11785.  
XX

PR 04-JUN-1997; 9705-0869309.  
PR 03-JUL-1996; 96US-0645601.  
PR 13-AUG-1996; 96US-0656381.  
XX  
PA (CHIR ) CHIRON CORP.  
XX  
PI Allen JR, Barber JR, Boder M, Chang SM, Chong K;  
PI De la Vega D, Depolo NJ, Greengard J, Hsu DC, Idanez CE;  
PI Jolly DJ, Mittlestaedt DM, Prussak CE, Respass JG;  
XX  
XX WPI; 1998-086967/08.  
DR P-PSDB; AAW4372.  
XX  
PT New replication defective recombinant retroviruses - which express B  
PT domain-deleted human factor VIII or human factor IX for the  
PT treatment of haemophilia  
XX  
XX  
PS Claim 6; Page 174-175; 236pp; English.  
XX  
CC This DNA sequence includes a coding region for the B domain  
CC deletion mutant SQN (see AAW4372) of human Factor VIII. The SQN  
CC mutant is created by fusing Ser-743 to Gln-1638 of native Factor  
CC VIII (see AAW4373) to form a Ser-Gln-Asn (SQN) link between the A2  
CC and A3 Factor VIII domains. When compared to plasmid-derived  
CC Factor VIII, the SQN deletion does not influence the in vivo  
CC pharmacokinetics, but the reduced size of the molecule appears to  
CC decrease proteolytic degradation. The invention relates to  
CC preparations of replication defective recombinant retrovirus (RV)  
CC expressing a B domain-deleted human Factor VIII protein, where the  
CC recombinant RV is capable of infecting human cells, is resistant  
CC to degradation by human complement and is capable of inducing  
CC long-term (at least 30 days and up to 6 months or longer  
CC post-injection) systemic expression of Factor VIII when  
CC administered to a haemophilia A patient.  
XX  
SQ Sequence 4832 BP; 1381 A; 1089 C; 1055 G; 1307 T; 0 other;

Query Match 90.7%; Score 4321.6; DB 19; Length 4832;  
Best Local Similarity 98.8%; Pred. No. 0;  
Matches 4379; Conservative 0; Mismatches 9; Indels 42; Gaps 1;

Qy 270 catcaaatgagatctccacactgctctctctctgctcttggcatttgcgatttgc 329  
Dh 71 catgcaaatgagatctccacactgctctctctctgctcttggcatttgcgatttgc 130  
Qy 330 caccagaagaatacactggtgctgagtgagactgcatggagactatgcaaatgact 389  
Dh 131 caccagaagaatacactggtgctgagtgagactgcatggagactatgcaaatgact 190  
Qy 390 cgtgtgagctgctctgtagcagcaagatttctctcagagtgccaaatctttccattcaa 449  
Dh 191 cgtgtgagctgctctgtagcagcaagatttctctcagagtgccaaatctttccattcaa 250  
Qy 450 caactcagctgctacaaaagaagactggttgaagattcaagttcaacttttcaacat 509  
Dh 251 caactcagctgctacaaaagaagactggttgaagattcaagttcaacttttcaacat 310  
Qy 510 cgttaagccaaggccaacctgagatggtctgctagtgctcaccatccagctgaggttta 569  
Dh 311 cgttaagccaaggccaacctgagatggtctgctagtgctcaccatccagctgaggttta 370  
Qy 570 tgatcacgtgtctattacacttaagaacatggtctccaccctgtagcttctatgctgtc 629  
Dh 371 tgatcacgtgtctattacacttaagaacatggtctccaccctgtagcttctatgctgtc 430  
Qy 630 tgggtatcctactcaggaagctctctgagggagctgtaataatgatatcagaccagttcaag 689  
Dh 431 tgggtatcctactcaggaagctctctgagggagctgtaataatgatatcagaccagttcaag 430  
Qy 690 ggaagaagaagatgataaagctctccctggtggaagcacaatataatgctcgagagctcct 749  
Dh 491 ggaagaagaagatgataaagctctccctggtggaagcacaatataatgctcgagagctcct 550



QY	750	gaagagaaatggtccaaatggtcccttgacccacgtgtcccttaactactcaatctttctca	809
Db	551	gaagaagaaatggtccaaatggtcccttgacccacgtgtcccttaactactcaatctttctca	610
QY	810	tgtggaaccttgtaaagaacctgaaatccaaaggtccatttgtagcccttaagtatagtgaaga	869
Db	611	tgtggaaccttgtaaagaacctgaaatccaaaggtccatttgtagcccttaagtatagtgaaga	670
QY	870	agggaagcttgccaaaggaagaacacagaccttgcacaanaattatactaattttgtcgt	929
Db	671	agggaagcttgccaaaggaagaacacagaccttgcacaanaattatactaattttgtcgt	730
QY	930	atttgaatgaagggaaagtgtgacctcagaagaacaagaactcttgaatgcagataggga	989
Db	731	atttgaatgaagggaaagtgtgacctcagaagaacaagaactcttgaatgcagataggga	790
QY	990	tgtctcacctgtctggagccttggtccataaatgcacaacgtcaatagtgttatgtaaacggtc	1041
Db	791	tgtctcacctgtctggagccttggtccataaatgcacaacgtcaatagtgttatgtaaacggtc	850
QY	1050	ctgtccaggtctgatttgatgagccacaggaatacagctcatttgcagatgtgatatggga	1101
Db	851	ctgtccaggtctgatttgatgagccacaggaatacagctcatttgcagatgtgatatggga	910
QY	1110	caccaccttcgaagtgtgcacctcaaatctccctccgaaggtccaaactcttctgttgaggaaca	1161
Db	911	caccaccttcgaagtgtgcacctcaaatctccctccgaaggtccaaactcttctgttgaggaaca	970
QY	1170	tcgcgaagcgctcccttgagaaatctgcgcaataaactcttcactacgtctccaaacactctgtat	1221
Db	971	tcgcgaagcgctcccttgagaaatctgcgcaataaactcttcactacgtctccaaacactctgtat	1031
QY	1230	ggaccttggagacgtttcctaactgtttgttcabatactctccacacaatgatgtgcattga	1281
Db	1031	ggaccttggagacgtttcctaactgtttgttcabatactctccacacaatgatgtgcattga	1091
QY	1290	agcttatgtccaagttagacaaacttccagaaggaaaccccaactcagaatgaanaataata	1341
Db	1091	agcttatgtccaagttagacaaacttccagaaggaaaccccaactcagaatgaanaataata	1151
QY	1350	agaagcgggaagaacatgatgatgtatctcttaactgtatcttcgaatgtgaatgtgtcaggttga	1401
Db	1151	agaagcgggaagaacatgatgatgtatctcttaactgtatcttcgaatgtgaatgtgtcaggttga	1211
QY	1410	tgaatgacaactctccttccttataccaaattgctcagttgccaagaagcatctctaaac	1461
Db	1211	tgaatgacaactctccttccttataccaaattgctcagttgccaagaagcatctctaaac	1271
QY	1470	cttggttaactatcaattgctgtctgtagaaggggaggttgagactatgagcccttaagctctgcg	1521
Db	1271	cttggttaactatcaattgctgtctgtagaaggggaggttgagactatgagcccttaagctctgcg	1331
QY	1530	ccccgaatgacagaaggtatataaagtccaatttgaanaacatgagccctcagcggaattgtag	1581
Db	1331	ccccgaatgacagaaggtatataaagtccaatttgaanaacatgagccctcagcggaattgtag	1391
QY	1590	gaagtacaaaaaagtccgatttatggtcatcacacagatgaaacctttaaagactcgtgaagc	1641
Db	1391	gaagtacaaaaaagtccgatttatggtcatcacacagatgaaacctttaaagactcgtgaagc	1451
QY	1650	tatcagaatgaataacagagaactctgggaaccttaacttatgggggaaggttgggaacacact	1701
Db	1451	tatcagaatgaataacagagaactctgggaaccttaacttatgggggaaggttgggaacacact	1511
QY	1710	gttgattatatttaagaatcagaagcaagacacatataaactataaccttaacctcaggaatcac	1761
Db	1511	gttgattatatttaagaatcagaagcaagacacatataaactataaccttaacctcaggaatcac	1571
QY	1770	tgaatgtccgtctcttgtcatccaagggagatccaacaaagtggttaaacatttgaagattt	1821
Db	1571	tgaatgtccgtctcttgtcatccaagggagatccaacaaagtggttaaacatttgaagattt	1631

QY	1830	tcattcttcgcagagaataattccaataataatgacacgaagtcgtgtgaagaatggcgc	1883
Dp	1631	tcacattctgcacagagaaatlatccaataataatgacagtgacgtgtgaagaatggcgc	1690
QY	1890	aactcaatcaaatccctcgtgcgcctgcagccgcctatgactagtttcgttaaatgtgagag	1949
Dp	1691	aactcaatcagaatccctcgtgcgcctgcagccgcctatgactagtttcgttaaatgtgagag	1750
QY	1950	agatctagcttcagagactcattggccctccctccatctctgtacaaagaatctgtatgaca	2009
Dp	1751	agactctagcttcagagactcattggccctccctccatctctgtacaaagaatctgtatgaca	1810
QY	2010	aaggggaaaccagataatgtgcagaaagagaaatgcataccgtttttctgtatttgatga	2068
Dp	1811	aaggggaaaccagataatgtgcagaaagagaaatgcataccgtttttctgtatttgatga	1870
QY	2070	gaaccggaagctgtgtaccttcacagaaataacacgctcttccccaatccagctgtgagat	2128
Dp	1871	gaaccggaagctgtgtaccttcacagaaataacacgctcttccccaatccagctgtgagat	1930
QY	2130	gcgaacttgagataatccaaagattccaagcttccaaacatcatgacagataatgagctatgt	2188
Dp	1931	gcgaacttgagataatccaaagattccaagcttccaaacatcatgacagataatgagctatgt	1990
QY	2190	tttctgatagttgcaggtctgtcagttttgtttgcatatgaggtgcatactgttatcctaa	2249
Dp	1991	tttctgatagttgcaggtctgtcagttttgtttgcatatgaggtgcagtaactgttatcctaa	2050
QY	2250	cattgtagacacagactgactctctctctgtctctctctctctgtatatataccttcaaacaca	2308
Dp	2051	cattgtagacacagactgactctctctctgtctctctctctctgtatatataccttcaaacaca	2110
QY	2310	aatgtctataagaacaacactccacccatcccatctccagagagaacagtccttcatgtc	2368
Dp	2111	aatgtctataagaacaacactccacccatcccatctccagagagaacagtccttcatgtc	2170
QY	2370	gatvgaaaacaccaggtctatgtgattctctgggtgcacaactcaagacttctggaaacagag	2428
Dp	2171	gatvgaaaacaccaggtctatgtgattctctgggtgcacaactcaagacttctggaaacagag	2230
QY	2430	catgacccgccttaattgaagttttctagtttgacaaagaacactggtgtatattacagaga	2489
Dp	2231	catgacccgccttaattgaagttttctagtttgacaaagaacactggtgtatattacagaga	2290
QY	2490	cagttatgaaagataattcagacatactgtctgagtaaaaaaatgcccattgaaaccaagag	2549
Dp	2291	cagttatgaaagataattcagacatactgtctgagtaaaaaaatgcccattgaaaccaagag	2350
QY	2550	cttctcccgaaattccaagacaacctgacatgacaaaagcaattaatgcccacccacc	2609
Dp	2351	cttctcccgaaattccaagacaacctgacatgacaaaagcaattaatgcccacccacc	2368
QY	2610	agctcttbaaacgcgcataacgcggagaaataactgcgtactactctctcagatcaagaagaga	2666
Dp	2369	agctcttbaaacgcgcataacgcggagaaataactgcgtactactctctcagatcaagaagaga	2428
QY	2670	aattgactatgatgatataccatcatcagttgaaatgaaagaagaatttgacattatgata	2729
Dp	2429	aattgactatgatgatataccatcatcagttgaaatgaaagaagaatttgacattatgata	2488
QY	2730	tgaagatgaanaatcaagagcccgcgagctttcaaaaagaanaacagacataattatgtc	2789
Dp	2489	tgaagatgaanaatcaagagcccgcgagctttcaaaaagaanaacagacataattatgtc	2548
QY	2790	tgcagtgagaaagctctgagataatgagataagtaagcttcccaacatgctctaaagaaag	2849
Dp	2549	tgcagtgagaaagctctgagataatgagataagtaagcttcccaacatgctctaaagaaag	2608
QY	2850	ggctccaaagtgcgaggtctccctcaagttcaagaagaagtgtgtttccagagaatttactcgtag	2909
Dp	2609	ggctccaaagtgcgaggtctccctcaagttcaagaagaagtgtgtttccagagaatttactcgtag	2668
QY	2910	ctctcttaaccagacccttataacqvtgaaanaacaaatgaaactttggagactctctggcc	2969



|||||  
Db 2669 ctcccttactcagccctataccgttgaggaactaaatgaaacatttggactccctgggccc 2728  
QY 2970 atataataagcgagaagtgttaagaataatcaatggttaacttccagaatacagcctctcg 3029  
Db 2729 atataataagcgagaagtgttaagaataatcaatggttaacttccagaatacagcctctcg 2788  
QY 3030 tccctatctctctctctctctctctctctctctctctctctctctctctctctctctctct 3089  
Db 2789 tccctatctctctctctctctctctctctctctctctctctctctctctctctctctctct 2848  
QY 3090 acctagaaaaaacttctgcaagcctaaatgaacccaacttacttctggaagtgcacaa 3149  
Db 2849 acctagaaaaaacttctgcaagcctaaatgaacccaacttacttctggaagtgcacaa 2908  
QY 3150 tcatatgcccccaactaaataatgagtttgactcgaaaagccttggtctattctctgagt 3209  
Db 2909 tcatatgcccccaactaaataatgagtttgactcgaaaagccttggtctattctctgagt 2968  
QY 3210 tgacctgaaaaaagatgtgcaactcagccctgattggaacccctctctgctcgcacactaa 3269  
Db 2969 tgacctgaaaaaagatgtgcaactcagccctgattggaacccctctctgctcgcacactaa 3028  
QY 3270 cacactgaaccctgctcactatggagagacagtgaacagtaacaggaatttgctctgtttcac 3329  
Db 3029 cacactgaaccctgctcactatggagagacagtgaacagtaacaggaatttgctctgtttcac 3088  
QY 3330 catcttttgatgagacccaaagcctgttacttcaactgaaataatggaagaactgcagagc 3389  
Db 3089 catcttttgatgagacccaaagcctgttacttcaactgaaataatggaagaactgcagagc 3148  
QY 3390 tccctgcaatatccagaatggaatgccacttttaagaagaatlatcgcttccatgtaac 3449  
Db 3149 tccctgcaatatccagaatggaatgccacttttaagaagaatlatcgcttccatgtaac 3208  
QY 3450 caatggtacataatggaatcactacacttggtctagtaattgctcagagatcaaaagattcg 3509  
Db 3209 caatggtacataatggaatcactacacttggtctagtaattgctcagagatcaaaagattcg 3268  
QY 3510 atggtatctgctcagcattggtgcagcaatggaataatccatctatctatcttaacttgatgaca 3569  
Db 3269 atggtatctgctcagcattggtgcagcaatggaataatccatctatctatcttaacttgatgaca 3328  
QY 3570 tgtgtcactgtaacgaataaagaagagataaaatggtgcactgtacacatctctacacag 3629  
Db 3329 tgtgtcactgtaacgaataaagaagagataaaatggtgcactgtacacatctctacacag 3388  
QY 3630 tgtttttgagacagtggaaatgttaacacatccaaagctggaatttggcgggtggaaatgcct 3689  
Db 3389 tgtttttgagacagtggaaatgttaacacatccaaagctggaatttggcgggtggaaatgcct 3448  
QY 3690 tatttgagagcatctacatgctctggtgagacacacttttctgtgtgtacagcaataatgt 3749  
Db 3449 tatttgagagcatctacatgctctggtgagacacacttttctgtgtgtacagcaataatgt 3508  
QY 3750 tccagactccccccttggaatggtctctgacacatctagaagatttccagattacagcttcag 3809  
Db 3509 tccagactccccccttggaatggtctctgacacatctagaagatttccagattacagcttcag 3568  
QY 3810 acaatatgagcagtggtgccccaaagctgtgcagactcattatctcggatcaatcaatgc 3869  
Db 3569 acaatatgagcagtggtgccccaaagctgtgcagactcattatctcggatcaatcaatgc 3828  
QY 3870 ctgtgagacacaaagagccttctctgtgacagagtgagatctgtgtgcacaaatgattat 3929  
Db 3629 ctgtgagacacaaagagccttctctgtgacagagtgagatctgtgtgcacaaatgattat 3688  
QY 3930 tcaagcatcaaaagacagaggtgcctgcagaaagttctccagcctcaactcctcaagtt 3989  
Db 3689 tcaagcatcaaaagacagaggtgcctgcagaaagttctccagcctcaactcctcaagtt 3748  
QY 3990 tatcatcatgtatagttctgtatgaggaagatggtgcagacttatcgagaatctccactgg 4049  
|||||

Db 3749 tatcatcatgtatagttctgtatgaggaagatggtgcagacttatcgagaatctccactgg 3808  
QY 4050 aaaccttaatggtcttctcttcttggaatgtgagttcatctgggataataaacacatattttaa 4109  
Db 3809 aaaccttaatggtcttctcttcttggaatgtgagttcatctgggataataaacacatattttaa 3868  
QY 4110 cccctcaatatgtctgtgatacatccgttgcaccccacacatataagcatctgcagacac 4169  
Db 3869 cccctcaatatgtctgtgatacatccgttgcaccccacacatataagcatctgcagacac 3928  
QY 4170 tcttcgcatggaatgtgagtggtctgtgataatgaatgltgcagatgcacatggaatgga 4229  
Db 3929 tcttcgcatggaatgtgagtggtctgtgataatgaatgltgcagatgcacatggaatgga 3988  
QY 4230 gaatgaagcaatatccagatgcagagattactgttcaacttcaacttcaacttcaacttcaact 4289  
Db 3989 gaatgaagcaatatccagatgcagagattactgttcaacttcaacttcaacttcaacttcaact 4048  
QY 4290 cacctgtctcctctcaaaaagctcgaactcaccctccaaaggaaggaatgacccggagac 4349  
Db 4049 cacctgtctcctctcaaaaagctcgaactcaccctccaaaggaaggaatgacccggagac 4108  
QY 4350 tcaagtgaaataatccaaaagatggtcgtgcagatggtgacttccagaagaacatgaagtac 4409  
Db 4109 tcaagtgaaataatccaaaagatggtcgtgcagatggtgacttccagaagaacatgaagtac 4168  
QY 4410 aggaagtaactcactccaggaataaactctcgtctaccagatgtatgtgaagagttcct 4469  
Db 4169 aggaagtaactcactccaggaataaactctcgtctaccagatgtatgtgaagagttcct 4228  
QY 4470 catctccagcagtcacagaatggtcactcagtgactctcttcttccagaatggaatgaataa 4529  
Db 4229 catctccagcagtcacagaatggtcactcagtgactctcttcttccagaatggaatgaataa 4288  
QY 4530 ggttttccaggaataatcaaacctccttccaacctgtgtgtgaactctctagaccacagtt 4589  
Db 4289 ggttttccaggaataatcaaacctccttccaacctgtgtgtgaactctctagaccacagtt 4248  
QY 4590 actgactcgttactccttcgaattaccccccaagattggtgtgcacccaattgacctgggat 4649  
Db 4349 actgactcgttactccttcgaattaccccccaagattggtgtgcacccaattgacctgggat 4408  
QY 4650 ggaagttctgtgctgcagagcacaggaacctctactgactgacagcagatgc 4699  
Db 4409 ggaagttctgtgctgcagagcacaggaacctctactgactgacagcagatgc 4458  
  
RESULT 9  
AAA49232  
ID AAA49232 standard; DNA; 12445 BP.  
XX  
AC AAA49232;  
XX  
DT 26-SEP-2000 (first entry)  
XX  
DE Vector HSReneo for transforming endothelial cells.  
XX  
KW Hemostatic; antianemic; antidiabetic; nootropic; neuroprotective; vector;  
KW osteopathic; antislacking; immunostimulant; gene therapy; collage;  
KW endothelial cell; peripheral blood; buffy coat cell; VEGF; PCR primer;  
KW vascular endothelial growth factor; bovine brain extract; haemophilia;  
KW Factor VIII; human; transgene; adenosine deaminase deficiency; ss;  
KW sickle cell anaemia; thalassemia; diabetes; alpha-antitrypsin deficiency;  
KW Alzheimer's disease; brain disease; heart disease; immune system defect;  
KW bone fracture; osteoporosis.  
XX  
OS Homo sapiens.  
OS Synthetic.  
PN WO200032750-A1.  
XX  
PD 08-JUN-2000.  
XX



PF 24-NOV-1999; 99WO-US28033.  
XX  
PR 24-NOV-1998; 98US-0109687.  
XX  
XX (MIND ) UNIV MINNESOTA.  
PA (UYEM-) UNIV EMORY.  
PA (HEBB/) HEBBEL R P.  
PA (LIN/) LIN Y.  
XX (LOLL/) LOLLAR J S.  
PI HEBBEL RP, Lin Y, LOLLAR JS.  
XX  
DR WPI: 2000-412303/35.  
XX  
PT Expanding population of endothelial cells useful to biocompatibilize  
PT implantable medical devices comprises contacting buffy coat cells with  
PT collagen I coated surface in culture medium comprising vascular  
PT endothelial growth factor -  
XX  
XX  
PS Claim 19; Fig 4; 53pp; English.  
XX  
CC The invention relates to a method for expanding the population of  
CC endothelial cells (EC) obtained from peripheral blood, by culturing,  
CC in contact with a collagen I coated surface, buffy coat cells obtained  
CC from peripheral mammalian blood in the presence of a culture medium  
CC containing vascular endothelial growth factor (VEGF) and free of bovine  
CC brain extract. EC are useful for treating hemophilia by introducing it  
CC into the blood stream of a mammal, so that an effective amount of  
CC Factor VIII protein is secreted in the blood stream of the mammal.  
CC This sequence represents the vector HSOReNeo where the human factor  
CC VIII gene into which an enhanced green fluorescent protein coding  
CC sequence (H8Q) has been inserted, is subcloned. Transgenic EC transduced  
CC in vitro are useful for improving prosthetic implants. EC is also useful  
CC for diagnosing clotting disorders where indication or disease is  
CC associated with a reduction in the activity of an enzyme. EC is also  
CC useful in gene therapy for treating the variety of diseases including  
CC adenosine deaminase deficiency, sickle cell anemia, thalassemia,  
CC hemophilia, diabetes, alpha-antitrypsin deficiency, brain disorders  
CC such as Alzheimer's disease, heart diseases, defects in immune system,  
CC for repairing bone fractures and to treat or prevent osteoporosis.  
XX  
XX  
SQ Sequence 12445 BP; 3255 A; 2989 C; 2951 G; 3250 T; 0 other;

Query Match 90.6%; Score 4319; DB 21; Length 12445;  
Best Local Similarity 98.9%; Pred. No. 0;  
Matches 4374; Conservative 0; Mismatches 5; Indels 42; Gaps 1;

QY 266 ccacacatgcacaaatagaagctctccacctgctcttctgtgctcttggatctgcttta 325  
DB 763 ccacacatgcacaaatagaagctctccacctgctcttctgtgctcttggatctgcttta 822  
QY 326 gtccacacagaagaatactacacctgtgagtggaactgtcaatggacacatgtcaaatg 385  
DB 823 gtgcacacagaagaatactacacctgtgagtggaactgtcaatggacacatgtcaaatg 882  
QY 386 atctcgggtgagctgtgtggaacgaagattctctctctagagtgccaaatctttccat 445  
DB 883 atctcgggtgagctgtgtggaacgaagattctctctctagagtgccaaatctttccat 942  
QY 446 tcaaacacctgctgtgtacaaaaaagactcgtttgttgaatcaagttcaaccttttca 505  
DB 943 tcaaacacctgctgtgtgtacaaaaaagactcgtttgttgaatcaagttcaaccttttca 1002  
QY 506 acatcgtgaagccaagccaagcctggaatggtgtcgtatgagtcctaccatcagagtgagg 565  
DB 1003 acatcgtgaagccaagccaagcctggaatggtgtcgtatgagtcctaccatcagagtgagg 1062  
QY 566 ttatattgaagtggtatcaacttaagaacatgagtcctccatccttcaagttctaatg 625  
DB 1063 ttatattgaagtggtatcaacttaagaacatgagtcctccatccttcaagttctaatg 1122  
QY 626 ctgtgtgtatcactactgtaagaagctctctgaggaagctgaatatgatacgaaccagtc 685

DB 1123 ctgtgtgtatcactactgtaagaagctctctgaggaagctgaatatgatacgaaccagtc 1182  
QY 686 aagagggaagaagaatgatataaagttctccctgtgtgaagccaacatgctctggagg 745  
DB 1183 aagagggaagaagaatgatataaagttctccctgtgtgaagccaacatgctctggagg 1242  
QY 746 tccgtgaagaagaatggttcccaatggcctctgacccacgtctgacctactactatcctt 805  
DB 1243 tccgtgaagaagaatggttcccaatggcctctgacccacgtctgacctactactatcctt 1302  
QY 806 ctcatgtgacctgtgtlaaaaaactgttaattcaagcctcaattggaagccctactagatgta 865  
DB 1303 ctcatgtgacctgtgtlaaaaaactgttaattcaagcctcaattggaagccctactagatgta 1362  
QY 866 gagaaggagctctgtccaaggaagaacagagctctgccaattatctactctttg 925  
DB 1363 gagaaggagctctgtccaaggaagaacagagctctgccaattatctactctttg 1422  
QY 926 ctgtatttgaatgaagggaagaagttgtgcactcagaacaagaactccttgatgcagata 985  
DB 1423 ctgtatttgaatgaagggaagaagttgtgcactcagaacaagaactccttgatgcagata 1482  
QY 986 gggatgtgcactgtgtctcggcctgtgcctaaatgcacacagtcatagtgtatgtlaaca 1045  
DB 1483 gggatgtgcactgtgtctcggcctgtgcctaaatgcacacagtcatagtgtatgtlaaca 1542  
QY 1046 ggtctctgcagagctgtatgtgatgtccacagaatacagtcatagtgtatgtlaaga 1105  
DB 1543 ggtctctgcagagctgtatgtgatgtccacagaatacagtcatagtgtatgtlaaga 1602  
QY 1106 tgggacacacactcctgaatgtgcactcaatattcctcgaaggltcacacattctgtgaaga 1165  
DB 1603 tgggacacacactcctgaatgtgcactcaatattcctcgaaggltcacacattctgtgaaga 1662  
QY 1166 accatgcgcagagctctctgtgaataatctgcgcaataaacttctctactgctccaacactct 1225  
DB 1663 accatgcgcagagctctctgtgaataatctgcgcaataaacttctctactgctccaacactct 1722  
QY 1226 tgatgacctgtggacagttctactgtttgttcatatctctccacacacatgataggca 1285  
DB 1723 tgatgacctgtggacagttctactgtttgttcatatctctccacacacatgataggca 1782  
QY 1286 tggaaagctatgttcaaaagtagacagctgtccagaagaagccccaactacgaatgaanaata 1345  
DB 1783 tggaaagctatgttcaaaagtagacagctgtccagaagaagccccaactacgaatgaanaata 1842  
QY 1346 atgaagaagcggaagactatgatataatcttactgatttgaanaatgagatgtgtgaagt 1405  
DB 1843 atgaagaagcggaagactatgatataatcttactgatttgaanaatgagatgtgtgaagt 1902  
QY 1406 ttgatgtgaacaactctctctccttlaaccaaatctgcctcagttgcgaagaagacatccta 1465  
DB 1903 ttgatgtgaacaactctctctccttlaaccaaatctgcctcagttgcgaagaagacatccta 1962  
QY 1466 aaacttgggtacatcaactatgtctgtcgtgaagaggaagctgtggactatgctcccttagtcc 1525  
DB 1963 aaacttgggtacatcaactatgtctgtcgtgaagaggaagctgtggactatgctcccttagtcc 2022  
QY 1526 tcggcccccagatgacagaagttataaagttcaattttgaacaatggccctcagcggaattg 1585  
DB 2023 tcggcccccagatgacagaagttataaagttcaattttgaacaatggccctcagcggaattg 2082  
QY 1586 gtgagaagtacaaaaaagctccgaatttattgtcatatacagaatgaacaccttaagacgtgt 1645  
DB 2083 gtgagaagtacaaaaaagctccgaatttattgtcatatacagaatgaacaccttaagacgtgt 2142  
QY 1646 aagcatatcagcatgaatcagaatctctggaccttacttatttgggaagtgtgagaca 1705  
DB 2143 aagcatatcagcatgaatcagaatctctggaccttacttatttgggaagtgtgagaca 2202  
QY 1706 caactgttataatttataagaatcaagcagacacatatataacttaccctccagga 1765



Dh 2203 cactgttgatataataaagaatcaagcaagcaacatataacatctaccctcagcaa 2262  
Qy 1766 tcaactgagctgcgtccctttrgattcaaggagatccaaaagtgtaaacattggaag 1895  
Dh 2263 tcaactgagctgcgtccctttrgattcaaggagatccaaaagtgtaaacattggaag 2332  
Qy 1826 atttccaattctgcagaggaataatcaataataatgagacagtagctgaaagatg 1885  
Dh 2233 atttccaattctgcagaggaataatcaataataatgagacagtagctgaaagatg 2382  
Qy 1886 ggccaactaaatcagatccctcgtgctgcagccgattactagtttcgttaatatg 1945  
Dh 2383 ggccaactaaatcagatccctcgtgctgcagccgattactagtttcgttaatatg 2442  
Qy 1946 agagaatctagcttccaggaatcttgccctccctccatctgtttaaagaatctag 2005  
Dh 2443 agagaatctagcttccaggaatcttgccctccctccatctgtttaaagaatctag 2502  
Qy 2006 atcaagaaggaacaacataatgctcagacaaggaatgtaacccgttttctgtattg 2065  
Dh 2503 atcaagaaggaacaacataatgctcagacaaggaatgtaacccgttttctgtattg 2562  
Qy 2066 atgagaacccgaagctggttaccctcagagaaatatacaagcttctcccaatccagctg 2125  
Dh 2563 atgagaacccgaagctggttaccctcagagaaatatacaagcttctcccaatccagctg 2622  
Qy 2126 gagtgcagcttgagatccagagttccaagctcccaatcatactgacacagcataatggtc 2185  
Dh 2623 gagtgcagcttgagatccagagttccaagctcccaatcatactgacacagcataatggtc 2682  
Qy 2186 atgtttctgtagtttgcagttgcaattgttctgcatagtgtagtaactgtagatctc 2245  
Dh 2683 atgtttctgtagtttgcagttgcaattgttctgcatagtgtagtaactgtagatctc 2742  
Qy 2246 taagcatatggagcaacagactgactctctctgtctctctctctgtagatatacctccaac 2305  
Dh 2743 taagcatatggagcaacagactgactctctctgtctctctctctgtagatatacctccaac 2802  
Qy 2306 acaaatggtctatgaagaacacatacctccatcccatctcagggagaacttctca 2365  
Dh 2803 acaaatggtctatgaagaacacatacctccatcccatctcagggagaacttctca 2862  
Qy 2366 tgtctgattggaaaaaccaggtctatgattctctggtgtgcacaactcagacttctggaaca 2425  
Dh 2863 tgtctgattggaaaaaccaggtctatgattctctggtgtgcacaactcagacttctggaaca 2922  
Qy 2426 gaggcattgacccgcttactgaaagttctctagttgtgacaagaacactggtgattatacg 2485  
Dh 2923 gaggcattgacccgcttactgaaagttctctagttgtgacaagaacactggtgattatacg 2982  
Qy 2486 aggcacattatgaagaatattcaagcatatctgtcgtgagttaaaaacaatgccaattgaacaa 2545  
Dh 2983 aggcacattatgaagaatattcaagcatatctgtcgtgagttaaaaacaatgccaattgaacaa 3042  
Qy 2546 gaaactctcccaaatccaagcacacccctagcactagcacaagcaatttaattgcacacc 2605  
Dh 3043 gaaactctcccaaatccaagcacacccctagcactagcacaagcaatttaattgcacacc 3060  
Qy 2606 caaccagttctgnaaagcccatcaacggygaataactcgttactactctccagtcagatcaag 2665  
Dh 3061 caaccagttctgnaaagcccatcaacggygaataactcgttactactctccagtcagatcaag 3120  
Qy 2666 aggaatattgactatgatataccatcatcagttgtaaatgagaagaagaatttgaacattc 2725  
Dh 3121 aggaatattgactatgatataccatcatcagttgtaaatgagaagaagaatttgaacattc 3180  
Qy 2726 atgattgagatgaataatcagaagcccgcaagcttccaagaagaacacgacacatttta 2785  
Dh 3181 atgattgagatgaataatcagaagcccgcaagcttccaagaagaacacgacacatttta 3240  
Qy 2786 ttgctgtagtgagagagctctggaattatagtgtagtaagtaagctcccaatgttctaagaa 2845  
Dh 3241 ttgctgtagtgagagagctctggaattatagtgtagtaagtaagtaagctcccaatgttctaagaa 3300

Qy 2846 acaagggtcccaagtggtgagctgtccctcagttccaagaagaattgtttccaagaatttactg 2905  
Dh 3301 acaagggtcccaagtggtgagctgtccctcagttccaagaagaattgtttccaagaatttactg 3360  
Qy 2906 atggtcccttactcaagcccttataccggtggaagactaaatgaacatttggactcctg 2965  
Dh 3361 atggtcccttactcaagcccttataccggtggaagactaaatgaacatttggactcctg 3420  
Qy 2966 ggccaatataagaagcaagaattgaaagataatcatctgtaactttcaagaatacagcct 3025  
Dh 3421 ggccaatataagaagcaagaattgaaagataatcatctgtaactttcaagaatacagcct 3480  
Qy 3026 ctggtccctattctcttattcttagcttatttttttgaagaagatccaaggagaagag 3085  
Dh 3481 ctggtccctattctcttattcttagcttatttttttgaagaagatccaaggagaagag 3540  
Qy 3086 cagaacctagaanaaactttgtcaagcctaaatgaacaaacttacttttgaagaatgctc 3145  
Dh 3541 cagaacctagaanaaactttgtcaagcctaaatgaacaaacttacttttgaagaatgctc 3600  
Qy 3146 aacatcatatgagcaacacataaagatgagtttgaactgcaaaagcctgggtatttctctg 3205  
Dh 3601 aacatcatatgagcaacacataaagatgagtttgaactgcaaaagcctgggtatttctctg 3660  
Qy 3206 atgttgaccttggaanaagatgtgcaatccagcctgatttggaacccctctgctgcacaa 3265  
Dh 3661 atgttgaccttggaanaagatgtgcaatccagcctgatttggaacccctctgctgcacaa 3720  
Qy 3266 ctacaacactgaacccctgctatgtggaacagtgagacatgcaagaatttctcgtctt 3325  
Dh 3721 ctacaacactgaacccctgctatgtggaacagtgagacatgcaagaatttctcgtctt 3780  
Qy 3326 tcaccatctttagatgagacccaagaagctggttacttcaactgaaataatggaagaacatgca 3385  
Dh 3781 tcaccatctttagatgagacccaagaagctggttacttcaactgaaataatggaagaacatgca 3840  
Qy 3386 gggtccctggaatataccagatggaagaatcccaattttaaagaagaattatgccttcattg 3445  
Dh 3841 gggtccctggaatataccagatggaagaatcccaattttaaagaagaattatgccttcattg 3900  
Qy 3446 caatcaatggtctataatagatgatacactaccctggtctagttaatggtctcagagatcaagaa 3505  
Dh 3901 caatcaatggtctataatagatgatacactaccctggtctagttaatggtctcagagatcaagaa 3960  
Qy 3506 ttcgatggtatctgtctcaagcattggtgcagcaatgaaacatccattcatcttcaagtg 3565  
Dh 3961 ttcgatggtatctgtctcaagcattggtgcagcaatgaaacatccattcatcttcaagtg 4020  
Qy 3566 gacatggttctcagctgtaacgaaaaaagaagatgtaaaatggtgacgttacaatctctatc 3625  
Dh 4021 gacatggttctcagctgtaacgaaaaaagaagatgtaaaatggtgacgttacaatctctatc 4080  
Qy 3626 cagggtcttttgaagaacgtgnaaatgttaccatccaagctggaatttggcgggtggaat 3685  
Dh 4081 cagggtcttttgaagaacgtgnaaatgttaccatccaagctggaatttggcgggtggaat 4140  
Qy 3686 gacttattggcagcatctatagcttgggaatgagcacacttttctggtgtacagcaata 3745  
Dh 4141 gacttattggcagcatctatagcttgggaatgagcacacttttctggtgtacagcaata 4200  
Qy 3746 agtgtcagacccctctgggaatgagcttctggaacatgaagaatttcagattacaact 3805  
Dh 4201 agtgtcagacccctctgggaatgagcttctggaacatgaagaatttcagattacaact 4260  
Qy 3806 caggaacataatgagcaatggtggcccaagcttggaacacttcatatttccgagatcaatca 3865  
Dh 4261 caggaacataatgagcaatggtggcccaagcttggaacacttcatatttccgagatcaatca 4320  
Qy 3866 atgcttgaggaaccaagaagaccttctctggaatcaagaatgagatctgtgacacaaatga 3925  
Dh 4321 atgcttgaggaaccaagaagaccttctctggaatcaagaatgagatctgtgacacaaatga 4380



QY	3926	tttttcaaggagatcaaaacccagggtgcccgtcaagaattcttccagcttatactctc	3985
Db	4381	ttattcaaggatcaagaaccacagggtgcctgcgtcaagaattcttccagcttatactctc	4440
QY	3986	agttatcatcatgttatagttctttatgttggaagaagtgtgcagacttalcagagaaattcca	4045
Db	4441	agttatcatcatgttatagttctttatgttggaagaagtgtgcagacttalcagagaaattcca	4500
QY	4046	ctggaaccttaatgtgtctctcttttgcaaatgttgatccaatctggataaacaacaatttt	4105
Db	4501	ctggaaaccttaatgtgtctctcttttgcaaatgttgatccaatctggataaacaacaatttt	4560
QY	4106	ttaaccccccaattatgtgtctgcgtacatactcgtttgcaacccaactcaatlaagattcgca	4165
Db	4561	ttaaccccccaattatgtgtctgcgtacatactcgtttgcaacccaactcaatlaagattcgca	4620
QY	4166	gcaactctcgaatbgagtttgaatggcctgtgatttaaatagtctgaacatgccattggaa	4225
Db	4621	gcaactctcgaatbgagtttgaatggcctgtgatttaaatagtctgaacatgccattggaa	4680
QY	4226	tggagagtaaaagcaatatcagatgcacagattactgtcttaccctacttaccataatgt	4285
Db	4681	tggagagtaaaagcaatatcagatgcacagattactgtcttaccctacttaccataatgt	4740
QY	4286	ttggcaactgtgtctcccttcaaaagcttgacttaccctcccaaggagagtaatgtcctgga	4345
Db	4741	ttggcaactgtgtctcccttcaaaagcttgacttaccctcccaaggagagtaatgtcctgga	4800
QY	4346	gaactcaggttaataatccaaagaagtgcgtctgcaagtgcgtcccaagaagacaataagaag	4405
Db	4801	gaactcaggttgaataataatccaaagaagtgcgtctgcaagtgcgtcccaagaagacaataagaag	4860
QY	4406	tcacagaagtaactactctcagggagtaaaatctctgtcttaccagcatgtatgtgaaggagt	4465
Db	4861	tcacagaagtaactactctcagggagtaaaatctctgtcttaccagcatgtatgtgaaggagt	4920
QY	4466	ttcccatcttcacgaagtcacagaatgagccataatgtggaactcctttttcagaatgycagaag	4525
Db	4921	ttcccatcttcacgaagtcacagaatgagccataatgtggaactcctttttcagaatgycagaag	4980
QY	4526	taaaaggttttcagggagaatcaagaactccttccacaactgtgtggaactctctagaccac	4585
Db	4981	taaaaggttttcagggagaatcaagaactccttccacaactgtgtggaactctctagaccac	5040
QY	4586	cgttactgactcgtctactcttcgaattlcaaccocagagttgtgtgcacacagattgcccctga	4645
Db	5041	cgttactgactcgtctactcttcgaattlcaaccocagagttgtgtgcacacagattgcccctga	5100
QY	4646	ggatgtgaggtgtctggtctcgaggcacaggaactctactga 4686	
Db	5101	ggatgtgaggtgtctggtctcgaggcacaggaactctactga 5141	
RESULT 10			
AADD00121			
ID	AAD00121	standard; DNA; 11933 BP.	
XX	AAD00121:		
DT	31-JUL-2000	(first entry)	
XX			
DE		Recombinant adeno associated vector construct, PAAV-F8-1.	
XX			
KM		Recombinant Adeno Associated Vector: PAAV; PAAV-F8-1; human Factor VIII;	
KW		hVFIII; hNF-3 albumin promoter; human elongation factor-1alpha; Eftalpa	
KW		human growth hormone; hGH; Inverted terminal repeat; ITR; haemophilina;	
KW		gene therapy; ds.	
OS		Adeno associated virus.	
XX			
FN	WO200023116-A1.		
XX			
DP	27-APR-2000		

PF	19-OCT-1999;	99WO-US24495.
XX		
XX	20-OCT-1998;	98US-0104994.
PR	24-MAR-1999;	99US-0125974.
PR	30-JUL-1999;	99US-0364862.
XX		
PA	(AVIG-) AVIGEN INC.	
XX		
P1	Couto LB, Colosi PC;	
XX		
DR	WPI, 2000-339536/29.	
PT		
PT	New recombinant adenovirus-associated vector, useful for gene therapy	
PT	to treat hemophilia, comprises at least a portion of Factor VIII	
PT	operably linked to control sequence -	
XX		
PS	Example 2; Fig 5; 92pp; English.	
XX		
CC	The present DNA sequence is a recombinant adenovirus-associated vector,	
CC	(rAAV) construct, PAAV-F8-1. This expression vector comprises the HNF-3	
CC	albumin promoter, the first intron (-573 to +985) of human elongation	
CC	factor-1alpha (EF1alpha) gene, human Factor VIII coding sequence (hFVIII)	
CC	and polyadenylation signal from human growth hormone (hGH). This sequence	
CC	is inserted between the AAV inverted terminal repeat (ITR) regions. The	
CC	hFVIII coding region comprises the heavy chain gene segment with the	
CC	first 57 bp encoding the 19 amino acid signal peptide, A1 and A2 domains	
CC	and 5 amino acids from the N-terminus of the B domain. The light chain	
CC	segment comprises the C-terminal 85 amino acids of B domain and the A3,	
CC	C1 and C2 domains. Both the heavy and light chain segments are cloned	
CC	into the same plasmid separated by 42 nucleotides coding for 14 residues	
CC	of the B domain, that is deleted. This plasmid is operably linked to	
CC	control sequences, that directs the transcription and translation of the	
CC	Factor VIII gene. The adeno-associated viral vectors are used for gene	
CC	therapy to treat hemophilia. This method allows prolonged expression of	
CC	therapeutic levels of Factor VIII in vivo. The rAAV are used for gene	
CC	therapy, because of their broad host range, safety profile and duration	
CC	of expression in the infected hosts.	
XX		
XX		
SQ	Sequence 11933 BP; 3258 A; 2818 C; 2717 G; 3140 T; 0 other;	
Query Match	89.4%; Score 4259.4; DB 21; Length 11933;	
Best Local Similarity	98.2%; Pred. No. 0;	
Matches 4348; Conservative	0; Mismatches	1; Indels 78; Gaps
OY	266 ccacatgccaatagagctccacacgtctctcttcgtgctcttgatctgcatc	325
DB	422 ccacatgccaatagagctccacacgtctctcttcgtgctcttgatctgcatc	481
OY	326 gtgcacaccagaagatactactctggtgagtgagtggaactgtcactgtggaataatgcaagt	385
DB	482 gtgcacaccagaagatactactctggtgagtgagtggaactgtcactgtggaataatgcaagt	541
OY	386 atctcgtgtagctgcctgtggaagcgaagattcctcctctagatggtccaaaatctttccat	445
DB	542 atctcgtgtagctgcctgtggaagcgaagattcctcctctagatggtccaaaatctttccat	601
OY	446 tcaacactcagtcgtgtgacaacaaaagactcgtttgtgataattcaaggttcaactttca	505
DB	602 tcaacactcagtcgtgtgacaacaaaagactcgtttgtgataattcaaggttcaactttca	661
OY	506 acatcgtcctaaagccaaagccacccctggaatggtgtcgtcgtgagtcctaccacccagctgag	565
DB	662 acatcgtcctaaagccaaagccacccctggaatggtgtcgtcgtgagtcctaccacccagctgag	721
OY	566 ttatgatacagtggtgtccttaccacttaagaacatggtctccatcctgtcagttcattg	625
DB	722 ttatgatacagtggtgtccttaccacttaagaacatggtctccatcctgtcagttcattg	781
OY	626 ctgtgtgtgtatcctactcgtgaagagctctctgagggagctggaatagatgatacgaacagtc	685
DB	782 ctgtgtgtgtatcctactcgtgaagagctctctgagggagctggaatagatgatacgaacagtc	841



QY 686 aagggagaaagatgataaagttccctgggtggaaagccatataatgtctggcagg 745  
|||||  
Db 842 aaaaaggaaagaaagataaagcttccctgggtggaagccatacatagctggcagg 901  
QY 746 tccctgaagaagatgtgtccaatgtgcctctgaaccacgtgtcccttaccatcatctt 805  
|||||  
Db 902 tccctgaagaagatgtgtccaatgtgcctctgaaccacgtgtcccttaccatcatctt 961  
QY 806 ctcatgtggaacctgttaaaagacttgaatccaggcccatgtggagccctactagta 865  
|||||  
Db 962 ctcatgtggaacctgttaaaagacttgaatccaggcccatgtggagccctactagta 1021  
QY 866 gagaaggagatctgtgccaagaagaaagacacagacctgtgccaattatatactattttg 925  
|||||  
Db 1022 gagaaggagatctgtgccaagaagaaagacacagacctgtgccaattatatactattttg 1081  
QY 926 ctgtatttgaatgaagggaagaaagttgtgacctcagaacaaagaacctcttgatgagagata 985  
|||||  
Db 1082 ctgtatttgaatgaagggaagaaagttgtgacctcagaacaaagaacctcttgatgagagata 1141  
QY 986 gggatgtgcatctgtctcggtgctgtgccttaaatgtcacacagtcaatgtgtataaaca 1045  
|||||  
Db 1142 gggatgtgcatctgtctcggtgctgtgccttaaatgtcacacagtcaatgtgtataaaca 1201  
QY 1046 ggtctgtccaggtctgattggaatgtgcacagaaatcagttatgtgcatgtatgaa 1105  
|||||  
Db 1202 ggtctgtccaggtctgattggaatgtgcacagaaatcagttatgtgcatgtatgaa 1261  
QY 1106 tgggacacacctctgtgaagtgtcacatcaatctcctcgaagggtcaacactctctgtgagga 1165  
|||||  
Db 1262 tgggacacacctctgtgaagtgtcacatcaatctcctcgaagggtcaacactctctgtgagga 1321  
QY 1166 accatgcgcagggctctgtgaaatctgcgcaataactcttccctactgtctcaaacactct 1225  
|||||  
Db 1322 accatgcgcagggctctgtgaaatctgcgcaataactcttccctactgtctcaaacactct 1381  
QY 1226 tgaatggaccttggagcaatcttctaagtctgtctatctcttcccaacaacatgagtga 1285  
|||||  
Db 1382 tgaatggaccttggagcaatcttctaagtctgtctatctcttcccaacaacatgagtga 1441  
QY 1286 tgggaagttaatgtcaaaagttagacagcgtctcagaggaagaaaccacactcgaatgaanaata 1345  
|||||  
Db 1442 tgggaagttaatgtcaaaagttagacagcgtctcagaggaagaaaccacactcgaatgaanaata 1501  
QY 1346 atgaaagaagcgggaagactatgatagtactgtactgtacttgaatgtgagtgcaggt 1405  
|||||  
Db 1502 atgaaagaagcgggaagactatgatagtactgtactgtacttgaatgtgagtgcaggt 1561  
QY 1406 ttgagtgtgcaactctccttcccttcttcaaatctgtcgaagtgtgccaagaagacatcta 1465  
|||||  
Db 1562 ttgagtgtgcaactctccttcccttcttcaaatctgtcgaagtgtgccaagaagacatcta 1621  
QY 1466 aaacttgggtacatctacatgtgtctgaaagaggagactgtggaactatgtcccttaagtc 1525  
|||||  
Db 1622 aaacttgggtacatctacatgtgtctgaaagaggagactgtggaactatgtcccttaagtc 1681  
QY 1526 tgcgccccgatgacagaagtataaaagtcaatatcttgaacaatgtgcctcagcgagattg 1585  
|||||  
Db 1682 tgcgccccgatgacagaagtataaaagtcaatatcttgaacaatgtgcctcagcgagattg 1741  
QY 1586 gtaagaagtacaanaaagctcggatttatgtgcatacacagatgaaacctttaaagctgtg 1645  
|||||  
Db 1742 gtaagaagtacaanaaagctcggatttatgtgcatacacagatgaaacctttaaagctgtg 1801  
QY 1646 aagctatctcagcatgaaatcaggaatcttgggaaccttacttacttgaatgtggaaagtgtgagaca 1705  
|||||  
Db 1802 aagctatctcagcatgaaatcaggaatcttgggaaccttacttacttgaatgtggaaagtgtgagaca 1861  
QY 1706 cactgtgtatataatttaagaatacaagaagcagacataataacatctaccctcagga 1765  
|||||  
Db 1862 cactgtgtatataatttaagaatacaagaagcagacataataacatctaccctcagga 1921

QY 1766 tcaactgattccgtctcttgttatccaaggagatbaccaaaaggtgttaaaacatttgaag 1825  
|||||  
Db 1922 tcaactgattccgtctcttgttatccaaggagatbaccaaaaggtgttaaaacatttgaag 1981  
QY 1826 atttccaattcttgcagaggaataatccaataataaagtgaacagtgtactgtagaagt 1885  
|||||  
Db 1982 atttccaattcttgcagaggaataatccaataataaagtgaacagtgtactgtagaagt 2041  
QY 1886 ggcacaataatacagatccctcggtgtcctgtaaccgctatactactagtcttgcataatgtg 1945  
|||||  
Db 2042 ggcacaataatacagatccctcggtgtcctgtaaccgctatactactactagtcttgcataatgtg 2101  
QY 1946 agagagatctcagcttcaaggctcatgtgcctctccctcactctcgtcaaaaagatctgtag 2005  
|||||  
Db 2102 agagagatctcagcttcaaggctcatgtgcctctccctcactctcgtcaaaaagatctgtag 2161  
QY 2006 atcaaaaggagaacacagataatgtccaagagaaggaaatgtcatcctgttttctgtattt 2065  
|||||  
Db 2162 atcaaaaggagaacacagataatgtccaagagaaggaaatgtcatcctgttttctgtattt 2221  
QY 2066 atgagaacccgaagctgtgtactccaagagaaatatacaacgcttcttcccaatccagctg 2125  
|||||  
Db 2222 atgagaacccgaagctgtgtactccaagagaaatatacaacgcttcttcccaatccagctg 2281  
QY 2126 gagtgcagcttggaggatccagagttccaagcctccaacatcatgtccaagcatcaatgtgt 2185  
|||||  
Db 2282 gagtgcagcttggaggatccagagttccaagcctccaacatcatgtccaagcatcaatgtgt 2341  
QY 2186 atgttcttgaatgttgcagttgtcagttgtgttgcatagtgagttggaactgtgaactc 2245  
|||||  
Db 2242 atgttcttgaatgttgcagttgtcagttgtgttgcatagtgagttggaactgtgaactc 2401  
QY 2246 taagcatctggagcaacagactgactcttctgtcttcttctctctgtgatataccttcaac 2305  
|||||  
Db 2402 taagcatctggagcaacagactgactcttctgtcttcttctctctgtgatataccttcaac 2461  
QY 2306 acaaaaatgtgtctatgaagagacacactcaccatctccatctcgaaggaaacgttctca 2365  
|||||  
Db 2462 acaaaaatgtgtctatgaagagacacactcaccatctccatctcgaaggaaacgttctca 2521  
QY 2366 tgtcgaatgaaaaaaccaggtctatgtgattctgtgggtgcacaactcagaaatcttcgagaca 2425  
|||||  
Db 2522 tgtcgaatgaaaaaaccaggtctatgtgattctgtgggtgcacaactcagaaatcttcgagaca 2581  
QY 2426 gaggcatagaccgcttacttgaaggtttctctagtgtgtgaacaagacactggtgatattacg 2485  
|||||  
Db 2582 gaggcatagaccgcttacttgaaggtttctctagtgtgtgaacaagacactggtgatattacg 2641  
QY 2486 aggaacgttataagagatatttcagcatactgtgtgagttaaaacaatgtccattgaaacaa 2545  
|||||  
Db 2642 aggaacgttataagagatatttcagcatactgtgtgagttaaaacaatgtccattgaaacaa 2701  
QY 2546 gaagcttctccgaattccaagacacccctagcactagggcaaaagcaattaatgtccacc 2605  
|||||  
Db 2702 gaagcttctccgaattccaagacacacccctagcactagggcaaaagcaattaatgtccacc 2709  
QY 2606 caccagcttctgaaagcccatcaacgggaaataactcgttactactcttccagtcagataag 2665  
|||||  
Db 2710 caccagcttctgaaagcccatcaacgggaaataactcgttactactcttccagtcagataag 2743  
QY 2666 aggaatctgactatgatatcatcatcagttgaatgtgaagaggaagatttggacatt 2725  
|||||  
Db 2744 aggaatctgactatgatatcatcatcagttgaatgtgaagaggaagatttggacatt 2803  
QY 2726 atgattgaggaatgaanaatcagagcccccagcgttccaanaagaaacacgacactatttta 2785  
|||||  
Db 2804 atgattgaggaatgaanaatcagagcccccagcgttccaanaagaaacacgacactatttta 2863  
QY 2786 ttgctgcaggtgagagagcttctgtgattatgtgagttgagttccccaatgttctaagaa 2845  
|||||  
Db 2864 ttgctgcaggtgagagagcttctgtgattatgtgagttgagttccccaatgttctaagaa 2923  
QY 2846 acagggctcagaggtgcaggtctccctcagttcaagaaagtgttcttccaggaatttaccig 2905



Db 2924 acagggctcagagatgagatgtccctcagttccagaaagtgtttccaggaatttactg 2983  
 QY 2906 atggtctcttaactcagcccttataccgttgagaaactaaatgaacatttggactccgg 2965  
 Db 2984 atggtctcttactcagcccttataccgttgagaaactaaatgaacatttggactccgg 3043  
 QY 2966 ggcacataataagaagcagaatltgaagaataatacgtgaacttcagaaatcagacct 3025  
 Db 3044 ggcacataataagaagcagaatltgaagaataatacgtgaacttcagaaatcagacct 3103  
 QY 3026 ctggtccctatctcttcttattttagccttatttcttagaagaagatcagaaggcagaag 3085  
 Db 3104 ctggtccctatctcttcttattttagccttatttcttagaagaagatcagaaggcagaag 3163  
 QY 3086 cagaacctagaanaaactctgccaagcctaaatgaaccaaacttactcttggaaatgac 3145  
 Db 3164 cagaacctagaanaaactctgccaagcctaaatgaaccaaacttactcttggaaatgac 3223  
 QY 3146 aacatcataatgacccaactaaagaatgagttgactgcaagcctggccttattctctg 3205  
 Db 3224 aacatcataatgacccaactaaagaatgagttgactgcaagcctggccttattctctg 3283  
 QY 3206 atgtttgacctggaaaagaatgtgacactcagagcctgattggagcccttctgtctgcaca 3265  
 Db 3284 atgtttgacctggaaaagaatgtgacactcagagcctgattggagcccttctgtctgcaca 3343  
 QY 3266 ctacaacactggaacccctgctcatgtgagacagatgacagctcagaaatttgcctgctt 3325  
 Db 3344 ctacaacactggaacccctgctcatgtgagacagatgacagctcagaaatttgcctgctt 3403  
 QY 3326 tcacacacttctgtatgagaccaaagctgttactcactgaataataatgaaagaactgca 3385  
 Db 3404 tcacacacttctgtatgagaccaaagctgttactcactgaataataatgaaagaactgca 3463  
 QY 3386 gggtccctgcaataatcagaatgagaatccacttttaagagaattatctgtctccatg 3445  
 Db 3464 gggtccctgcaataatcagaatgagaatccacttttaagagaattatctgtctccatg 3523  
 QY 3446 caatcaactgtctacataatgatacaactcgtcttagtaatgtctcagatcaagaatga 3505  
 Db 3524 caatcaactgtctacataatgatacaactcgtcttagtaatgtctcagatcaagaatga 3583  
 QY 3506 ttcgaatgtatctgtctcagaatgagcagaatgaatgaatccatctctatctcactg 3565  
 Db 3584 ttcgaatgtatctgtctcagaatgagcagaatgaatgaatccatctctatctcactg 3643  
 QY 3566 gacatgtgttcaactgttcgaaaaaagagagtaaaatggcactgtgacaactctctac 3625  
 Db 3644 gacatgtgttcaactgttcgaaaaaagagagtaaaatggcactgtgacaactctctac 3703  
 QY 3626 caggtgttcttgagacagctggaatgttaccatccaaaagctggaatttggcgggtggaat 3685  
 Db 3704 caggtgttcttgagacagctggaatgttaccatccaaaagctggaatttggcgggtggaat 3763  
 QY 3686 gacctattgagcagacatctacatgctcgtggaatgagcaacttttctgtgttgcagcaata 3745  
 Db 3764 gacctattgagcagacatctacatgctcgtggaatgagcaacttttctgtgttgcagcaata 3823  
 QY 3746 agtgtcagaactcccttggaatgtctctgacacataggaattttagaatttagaatt 3805  
 Db 3824 agtgtcagaactcccttggaatgtctctgacacataggaattttagaatttagaatt 3883  
 QY 3806 caggaacataatgacagctggaaccccaaagctgacagacatcatatccgatatca 3865  
 Db 3884 caggaacataatgacagctggaaccccaaagctgacagacatcatatccgatatca 3943  
 QY 3866 atgccttgagacccaagagccttctcttgatcaagtgagatcgttggcacaatgca 3925  
 Db 3944 atgccttgagacccaagagccttctcttgatcaagtgagatcgttggcacaatgca 4003  
 QY 3926 ttattacggcatcaaaagcaggtgcccgtcagaagttctccagcctctacatctctc 3985

Db 4004 ttattacggcatcaaaagcaggtgcccgtcagaagttctccagcctctacatctctc 4063  
 QY 3986 agttatcatcatgatalagttcttgatggaaagatggcagacttaccagaatccaa 4045  
 Db 4064 agttatcatcatgatalagttcttgatggaaagatggcagacttaccagaatccaa 4123  
 QY 4046 ctggaaccttaatggtctcttcttggaatgtgattcaatcttggaataaacaataattc 4105  
 Db 4124 ctggaaccttaatggtctcttcttggaatgtgattcaatcttggaataaacaataattc 4183  
 QY 4106 ttaaccctccaattatgtctgatacatccgtttggacccaactcatatagatctgca 4165  
 Db 4184 ttaaccctccaattatgtctgatacatccgtttggacccaactcatatagatctgca 4243  
 QY 4166 gaactcttcgaatgagttgataaggctgtgatttaaatagttgcagatgccattggaa 4225  
 Db 4244 gaactcttcgaatgagttgataaggctgtgatttaaatagttgcagatgccattggaa 4303  
 QY 4226 tggagagtaaaagcaataatcagatgacagatctactgtcttaacttaccataatgt 4285  
 Db 4304 tggagagtaaaagcaataatcagatgacagatctactgtcttaacttaccataatgt 4363  
 QY 4286 ttggcaactgttctccttcaaaaagctgcagcttcaactcccaaggagagtaatgacctgga 4345  
 Db 4364 ttggcaactgttctccttcaaaaagctgcagcttcaactcccaaggagagtaatgacctgga 4423  
 QY 4346 gacctcaagtgaataatccaaaagatgtgctgcaagtggacttccaaagaacaaatgaaag 4405  
 Db 4424 gacctcaagtgaataatccaaaagatgtgctgcaagtggacttccaaagaacaaatgaaag 4483  
 QY 4406 tcaacagagtaactactcaatgaagtaaaatctcgtctaccagatgtatgtgaagagt 4465  
 Db 4484 tcaacagagtaactactcaatgaagtaaaatctcgtctaccagatgtatgtgaagagt 4543  
 QY 4466 tccatcactccagcagatcaaatgagccatgaatgagctctcttcttcagaatggcagaag 4525  
 Db 4544 tccatcactccagcagatcaaatgagccatgaatgagctctcttcttcagaatggcagaag 4603  
 QY 4526 taaaggttttcaagggaatcaagactccttcaaacctgtgtgaactctctagaccac 4585  
 Db 4604 taaaggttttcaagggaatcaagactccttcaaacctgtgtgaactctctagaccac 4663  
 QY 4586 cgttactgactcgtctactcttgaattcaaccacgaatttgggtgacacgaatttgcctga 4645  
 Db 4664 cgttactgactcgtctactcttgaattcaaccacgaatttgggtgacacgaatttgcctga 4723  
 QY 4646 ggaatgaggtctggtcgtgcagagcaagagcctctactgactgag 4692  
 Db 4724 ggaatgaggtctggtcgtgcagagcaagagcctctactgactgag 4770

RESULT 11  
 AAN90654  
 ID AAN90654 standard; DNA; 4275 BP.  
 AC AAN90654;  
 XX  
 AC AAN90654;  
 XX  
 DT 26-JUN-1990 (first entry)  
 XX  
 DE DNA encoding 740 Arg-1649 Glu human Factor VIII.C.  
 XX  
 KW Human Factor VIII.C; Ad.RE.neo; 740 Arg-1649 Glu Factor VIII.C;  
 KW haemophilila A.  
 OS Homo saplens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1..4275  
 FT /\*tag= a  
 PN EP306968-A.  
 XX  
 PD 15-MAR-1989.



XX 09-SEP-1988; 88EP-014769.  
XX  
PR 08-APR-1988; 88JP-0085454.  
XX  
PA (KAGA ) CHEMO-SERO-THERAP.  
PA (TEIJ ) TEIJIN LTD.  
XX  
PI Sugiyama T, Masuda K, Tajima Y, Yonemura H;  
DR WPI: 1989-078467/11.  
DR P-PSDB: AAP91165.  
XX  
PT Prodn. of recombinant human Factor-VIII-C -  
PT using animal cells transformed with a vector contg. the gene for  
PT Factor VIII:C and a promoter  
PS Fig 1(1) - 1(13); : 32pp; English.  
XX  
XX When translated, Arg-740 of the carboxyl terminus of the H chain is  
CC directly bonded by a peptide bond to Glu-1649 of the amino terminus of  
CC L chain. It is used to transform animal cells so that they produce  
CC human Factor VIII:C. A prefd. expression vector is plasmid Ad.RE.neo.  
CC The expression vector has at least one promoter upstream of AAP90654.  
CC The transformants can constantly and continuously produce human Factor  
CC VIII:C in high yield on a commercial scale. The human Factor VIII:C so  
CC produced is considered to corresp. to the smallest species of active and  
CC intact Factor VIII:C molecules in the human blood plasma. It is useful  
CC for treating haemophilia A patients.  
SQ Sequence 4275 BP; 1245 A; 941 C; 945 G; 1144 T; 0 other;

Query Match 87.6%; Score 4173; DB 10; Length 4275;  
Best Local Similarity 98.0%; Pred. No. 0;  
Matches 4270; Conservative 0; Mismatches 5; Indels 84; Gaps 1;

QY 328 gccaccagaagatgactctgtggtgagctggaactgtcatgtgacatattgcaaatgt 387  
DB 1 gccaccagaagatgactctgtggtgagctggaactgtcatgtgacatattgcaaatgt 60  
QY 388 ctccgtgagctgctgtgtgagcgaagattcctcctagagtgccaaacatttccatcc 447  
DB 61 ctccgtgagctgctgtgtgagcgaagattcctcctagagtgccaaacatttccatcc 120  
QY 448 aacacctcagtcgtgtacaaaagactctgttctgtagaattcaaggttcaacctttcaac 507  
DB 121 aacacctcagtcgtgtacaaaagactctgttctgtagaattcaaggttcaacctttcaac 180  
QY 508 atcgctcaagcgaagcgccactgagtggtctgtgtagtgcctacatccatcagctgaggt 567  
DB 181 atcgctcaagcgaagcgccactgagtggtctgtgtagtgcctacatccatcagctgaggt 240  
QY 568 tatgatacagtgatcatcaacttaagaacatggtctccatcctgttcagcttctagct 627  
DB 241 tatgatacagtgatcatcaacttaagaacatggtctccatcctgttcagcttctagct 300  
QY 628 gtgtgtgtatcctcacttggaagcttctgagggagctgaaatgatatgatacagacagtc 687  
DB 301 gtgtgtgtatcctcacttggaagcttctgagggagctgaaatgatatgatacagacagtc 360  
QY 688 aagggaagaagatgataaagttctcctctgtgtggaagcacaataatgtctgagcagtc 747  
DB 361 aagggaagaagatgataaagttctcctctgtgtggaagcacaataatgtctgagcagtc 420  
QY 748 ctgaaagaagatggttccaaatggcctctgacccaactgtgcttactaccatacttct 807  
DB 421 ctgaaagaagatggttccaaatggcctctgacccaactgtgcttactaccatacttct 480  
QY 808 catgtgacacctgttaagaacttgaattcaggccatctggagccctactagtatag 867  
DB 481 catgtgacacctgttaagaacttgaattcaggccatctggagccctactagtatag 540

QY 868 gaaggagtcctggccaaggaagacacagacttgcacaatttatactactttgct 927  
DB 541 gaaggagtcctggccaaggaagacacacacacttgcacaatttatactactttgct 600  
QY 928 gtaattgataaggaagatttggcaactcagaacaaagaactccttgatgcagatag 987  
DB 601 gtaattgataaggaagatttggcaactcagaacaaagaactccttgatgcagatag 660  
QY 988 gatgtgcatactgtctcgggctgctctaaatgacacagtcgaatgttatgtaacagg 1047  
DB 661 gatgtgcatactgtctcgggctgctctaaatgacacagtcgaatgttatgtaacagg 720  
QY 1048 tctctgcaggtctgtattgatagcacaggaatcagtcattggcattgagtggaatg 1107  
DB 721 tctctgcaggtctgtattgatagcacaggaatcagtcattgagtggaatg 780  
QY 1108 ggcaccactccttgaatgacatcaataatctcctcgaaggtacacacatttcttgaggaac 1167  
DB 781 ggcaccactccttgaatgacatcaataatctcctcgaaggtacacacatttcttgaggaac 840  
QY 1168 catgcgaagcgctccttggaatctgcgcaataacttccctactgtcctaacaactctg 1227  
DB 841 catgcgaagcgctccttggaatctgcgcaataacttccctactgtcctaacaactctg 900  
QY 1228 atggaccttggagacttctactgttttgcataatcttcccaacaatgtagcatg 1287  
DB 901 atggaccttggagacttctactgttttgcataatcttcccaacaatgtagcatg 960  
QY 1288 gaagcttatgtcaagtagacagctgtgccagaggaaccccaactaagaaatgaaataat 1347  
DB 961 gaagcttatgtcaagtagacagctgtgccagaggaaccccaactaagaaataat 1020  
QY 1348 gaagaagcggaagacacatgatatgtatcttactgattcgtgaatgtagtgcaggtt 1407  
DB 1021 gaagaagcggaagacacatgatatgtatcttactgattcgtgaatgtagtgcaggtt 1080  
QY 1408 gatgatgacaactctcctccttccatccaaatcgctcagcttgccaagagacatccaa 1467  
DB 1081 gatgatgacaactctcctccttccatccaaatcgctcagcttgccaagagacatccaa 1140  
QY 1468 acttgggtacatcatatgtctgtctgaaagaggaagctggagactgtcccttagtcttc 1527  
DB 1141 acttgggtacatcatatgtctgtctgaaagaggaagctggagactgtcccttagtcttc 1200  
QY 1528 gcccccagatgacgaagatataaagttcaatatgtgaacaatgcccctcagcgattgt 1587  
DB 1201 gcccccagatgacgaagatataaagttcaatatgtgaacaatgcccctcagcgattgt 1260  
QY 1588 aggaagtacaanaagttccgatttattggtacacagatgaaacctttaagactgtgaa 1647  
DB 1261 aggaagtacaanaagttccgatttattggtacacagatgaaacctttaagactgtgaa 1320  
QY 1648 gctatccagcatgaaatcagaagatcttggagaccttacttattatggggaagtgtgagacaca 1707  
DB 1321 gctatccagcatgaaatcagaagatcttggagaccttacttattatggggaagtgtgagacaca 1380  
QY 1708 ctgtgtattatatttaagaatacgaagcagacacatatcaactcctcagcgatc 1767  
DB 1381 ctgtgtattatatttaagaatacgaagcagacacatatcaactcctcagcgatc 1440  
QY 1768 actgatgtccgtccttctgtattcaaggaattacaaaaggtgttaaacatttgaagat 1827  
DB 1441 actgatgtccgtccttctgtattcaaggaattacaaaaggtgttaaacatttgaagat 1500  
QY 1828 ttcccaattctgcagaggaatatcaataataatgagcagtgagctgtagaagatg 1887  
DB 1501 ttcccaattctgcagaggaatatcaataataatgagcagtgagctgtagaagatg 1560  
QY 1888 ccaactcaaatcagatcctcgtgtgcgcgacccgtatattactagtgttctgtaatatggag 1947  
DB 1561 ccaactcaaatcagatcctcgtgtgcgcgacccgtatattactagtgttctgtaatatggag 1620  
QY 1948 agagatctagcttcagactcatgtgcccctcctcatctgctacaagaatctgtagat 2007



|||||  
Db 1621 agagactcgtccacgagcaccattgcccctccctcactcgtcacaagaatctgtagat 1660  
Oy 2008 caaagggaacacagataatgctcagacaagaagatgcaatcctgctttctgtattgtag 2067  
Db 1681 caaagggaacacagataatgctcagacaagaagatgcaatcctgctttctgtattgtag 1740  
Oy 2068 gaggaaacgaagctgtgtaacctcacagagaatalacaacgctttctcccaatccagctgga 2127  
Db 1741 gaggaaacgaagctgtgtaacctcacagagaatalacaacgctttctcccaatccagctgga 1800  
Oy 2128 gtgcacacttgagagatccagagctcccaagcctccaactcaatgctgcacaatcaatggcttat 2187  
Db 1801 gtgcacacttgagagatccagagctcccaagcctccaactcaatgctgcacaatcaatggcttat 1860  
Oy 2188 gttttgtagatgttgcagttgtcagttgtgttgcataaggtgagatactgtagatccta 2247  
Db 1861 gttttgtagatgttgcagttgtcagttgtgttgcataaggtgagatactgtagatccta 1920  
Oy 2248 agcattgagacacagactgtaactcctctctgtctctctctctgtgataaacctccaacac 2307  
Db 1921 agcattgagacacagactgtaactcctctctgtctctctctctgtgataaacctccaacac 1980  
Oy 2308 aaaaatgctctatgaagaacaacactcaacctatcccatctccagagagaactgcttcatg 2367  
Db 1981 aaaaatgctctatgaagaacaacactcaacctatcccatctccagagagaactgcttcatg 2040  
Oy 2368 tcgattggaanaacccaggtctctatgattctgtggtgacacacactcagacttcgcgaacaga 2427  
Db 2041 tcgattggaanaacccaggtctctatgattctgtggtgacacacactcagacttcgcgaacaga 2100  
Oy 2428 ggcatagcagcgccttaactgaaagtttctagtctgtgacaagaacaatggtgatactacgag 2487  
Db 2101 ggcatagcagcgccttaactgaaagtttctagtctgtgacaagaacaatggtgatactacgag 2160  
Oy 2488 gacagttatggaagatatatttcagcataactgtctgtgataaaaaaactgcctatggaacaga 2547  
Db 2161 gacagttatggaagatatatttcagcataactgtctgtgataaaaaaactgcctatggaacaga 2220  
Oy 2548 agcttctccagaaattccaagacccctcagactcagcaaaaagcaatttaatgcccacca 2607  
Db 2221 ----- 2220  
Oy 2608 ccagttctgaaacgcacatcaacggaataactcgttaactcctctcagtcagatcaagag 2667  
Db 2221 -----gaaataactcgttaactcctctcagtcagatcaagag 2256  
Oy 2668 gaaattgactatgatataaccatacatcagttgaaatggaagaagagatttggaaatttat 2727  
Db 2257 gaaattgactatgatataaccatacatcagttgaaatggaagaagagatttggaaatttat 2316  
Oy 2728 gatgagatgaaatcagagcccccgagcttccaagaagaacaacagacactatttat 2787  
Db 2317 gatgagatgaaatcagagcccccgagcttccaagaagaacaacagacactatttat 2376  
Oy 2788 gctgcagatgagagagctctggaatlaatggaatgagtagctccocacatgttctaagaac 2847  
Db 2377 gctgcagatgagagagctctggaatlaatggaatgagtagctccocacatgttctaagaac 2436  
Oy 2848 agggctcagaatgagcagtgctccctcagttccaagaagaatgttttccaaggaattacgagat 2907  
Db 2437 agggctcagaatgagcagtgctccctcagttccaagaagaatgttttccaaggaattacgagat 2496  
Oy 2908 ggcctcttaactcagaccctataacgctgaggaactaaatgaaatttggaagactctggg 2967  
Db 2497 ggcctcttaactcagaccctataacgctgaggaactaaatgaaatttggaagactctggg 2556  
Oy 2968 ccataataaagacagaagtgaagaataataatacatggttaacttcagaaaaatcagagcctc 3027  
Db 2557 ccataataaagacagaagtgaagaataataatacatggttaacttcagaaaaatcagagcctc 2616  
Oy 3028 cgtccctattcctctatctatctagccttatcttctaaggaagaatcagaagcagaagca 3087  
|||||

Db 2617 cgtccctattcctctatctatctagccttatcttaatgaagaatcagaagcagaagca 2676  
Oy 3088 gaacctagaanaaaactctgtccaagcctaatacgaacccaacttactttgaaagtgc 3147  
Db 2677 gaacctagaanaaaactctgtccaagcctaatacgaacccaacttactttgaaagtgc 2736  
Oy 3148 calcatatgcaaccactaaagatgagttgtaactgcaaaagcctgggacttaattctctgat 3207  
Db 2737 calcatatgcaaccactaaagatgagttgtaactgcaaaagcctgggacttaattctctgat 2796  
Oy 3208 gtgacactggaanaaagatgagcactcagagcctgatttgaccctctcgtgtccacat 3267  
Db 2797 gtgacactggaanaaagatgagcactcagagcctgatttgaccctctcgtgtccacat 2856  
Oy 3268 aacacactgaacccctgtctcactgagacaagtgcacacgcttgtagccctctcgtgtccacat 3327  
Db 2857 aacacactgaacccctgtctcactgagacaagtgcacacgcttgtagccctctcgtgtccacat 2916  
Oy 3328 accacttgcagtgaagacaaaagctggtactccaactgaaataatgaaagaactgcag 3387  
Db 2917 accacttgcagtgaagacaaaagctggtactccaactgaaataatgaaagaactgcag 2976  
Oy 3388 gctccctgcacatccagatggaagatcccaactttaagaagatatcgtctccatga 3447  
Db 2977 gctccctgcacatccagatggaagatcccaactttaagaagatatcgtctccatga 3036  
Oy 3448 atcaatgctacataatgatacactacactggtctgaataatggtcagatcaaaagat 3507  
Db 3037 atcaatgctacataatgatacactacactggtctgaataatggtcagatcaaaagat 3096  
Oy 3508 cgaatgatactcgtcagacatgaggcagaatgaaacaatccaattcaattcaattgag 3567  
Db 3097 cgaatgatactcgtcagacatgaggcagaatgaaacaatccaattcaattcaattgag 3156  
Oy 3568 catgtgtcactgtgacgaaagaaagagatataaaatggcactgtgacaatcctatcca 3627  
Db 3157 catgtgtcactgtgacgaaagaaagagatataaaatggcactgtgacaatcctatcca 3216  
Oy 3628 ggtgtttctgagacagtggaatgtlaccatccaagcctggaatttgagcggtgagatgc 3687  
Db 3217 ggtgtttctgagacagtggaatgtlaccatccaagcctggaatttgagcggtgagatgc 3276  
Oy 3688 cttaattgagacactacatgcttggaatgagcaacacttttctgtgtacagacaataag 3747  
Db 3277 cttaattgagacactacatgcttggaatgagcaacacttttctgtgtacagacaataag 3336  
Oy 3748 tgtcagactccctgggaatgagcttctgacacattagaagattccaagattacaagcttca 3807  
Db 3337 tgtcagactccctgggaatgagcttctgacacattagaagattccaagattacaagcttca 3396  
Oy 3808 ggaacaataatgacatgagtgccccaagctgacagactcatattccgatatcaatcaat 3867  
Db 3397 ggaacaataatgacatgagtgccccaagctgacagactcatattccgatatcaatcaat 3456  
Oy 3868 gctcagagaccaaagagcccttctctggaatcaaggtggaatcgtltgacacaatgagat 3927  
Db 3457 gctcagagaccaaagagcccttctctggaatcaaggtggaatcgtltgacacaatgagat 3516  
Oy 3928 attcaacgacatcaagacccaggtgtcccgctcagaagttcccagccctctacacatctccag 3987  
Db 3517 attcaacgacatcaagacccaggtgtcccgctcagaagttcccagccctctacacatctccag 3576  
Oy 3988 tttaatacatgataatctctgtagtgaagaagtggaagactataatcgaagaatccact 4047  
Db 3577 tttaatacatgataatctctgtagtgaagaagtggaagactataatcgaagaatccact 3636  
Oy 4048 ggaacctaaatgctccttctggaatgagatcctggaataaacaataatctt 4107  
Db 3637 ggaacctaaatgctccttctggaatgagatcctggaataaacaataatctt 3696  
Oy 4108 aacccccaattatgtgataatcaccgtttgcaccaactatcatatggaatttcgagc 4167  
Db 3697 aacccccaattatgtcgaatacatccgtttgcaccaactatcatatggaatttcgagc 3756



QY 4168 acctctgcagtgagatgagctgctgatttaattagttgcagcatgccaatggaatg 4227  
|||||  
Db 3757 acctctgcagtgagatgagctgctgatttaattagttgcagcatgccaatggaatg 3816  
QY 4228 gagagtaaaagcaatatacatgacagatgactgcttcattccttacttaccataatgtt 4287  
|||||  
Db 3817 gagagtaaaagcaatatacatgacagatgactgcttcattccttacttaccataatgtt 3876  
QY 4288 gccacccgtgtcctcttcaaaaagctgcagcttcaacctccaagggagagtaatgcttgaga 4347  
|||||  
Db 3877 gccacccgtgtcctcttcaaaaagctgcagcttcaacctccaagggagagtaatgcttgaga 3936  
QY 4348 cctcaggtgtaataccaataaagtgctgcaagtggaactgcagaagacaataaagtc 4407  
|||||  
Db 3937 cctcaggtgtaataccaataaagtgctgcaagtggaactgcagaagacaataaagtc 3996  
QY 4408 acaagagtaactactcagggagtaaaatctcctgtcttaccagcatgtaagtgaagagttc 4467  
|||||  
Db 3997 acaagagtaactactcagggagtaaaatctcctgtcttaccagcatgtaagtgaagagttc 4056  
QY 4468 ctctctccagcagtaagaatgagccatcagtgactctcttttttcagaatggcaagta 4527  
|||||  
Db 4057 ctctctccagcagtaagaatgagccatcagtgactctcttttttcagaatggcaagta 4116  
QY 4528 aaggttttcagggaaatcaagactccttcaacactgtgtgaaactcttagaccacag 4587  
|||||  
Db 4117 aaggttttcagggaaatcaagactccttcaacactgtgtgaaactcttagaccacag 4176  
QY 4588 ttactgactgcctactcttcgaatcacccccaagatgtgggtgcacagatgtcccttagg 4647  
|||||  
Db 4177 ttactgactgcctactcttcgaatcacccccaagatgtgggtgcacagatgtcccttagg 4236  
QY 4648 atggaagttctggctgcgagggcacaaggaaccttactga 4686  
|||||  
Db 4237 atggaagttctggctgcgagggcacaaggaaccttactga 4275  
  
RESULT 12  
AAN80446  
ID AAN80446 standard; DNA: 4275 BP.  
XX  
AC AAN80446;  
XX  
DT 10-OCT-1990 (first entry)  
XX  
DE Modified factor VIII:C sequence with the R740-E1649 deletion.  
XX  
KM Modified factor VIII:C; haemophilia; procoagulant;  
XX  
KM blood coagulation; RE deletion; ss.  
XX  
OS Homo sapiens.  
XX  
PN NC0800831-A.  
XX  
PD 11-FEB-1988.  
XX  
PF 31-JUL-1987; 87MO-US01814.  
XX  
PR 01-AUG-1986; 86US-0893375.  
XX  
PA (BIOJ ) BIOGEN NV.  
XX  
PI Pasek MP.  
XX  
DR WPI; 1988-049866/07.  
XX  
DR P-PSDB; AAP80267.  
XX  
PT New DNA sequences encoding modified factor VIII:C - with deletion of DNA  
XX encoding maturation polypeptide, useful for high yield transformation.  
XX  
PS Claim 3; Page 44-45-46-47; 97P; English.  
XX

CC The entire sequence encoding the maturation polypeptide of  
CC factor VIII:C is deleted, i.e. Arg 740-Glu 1649.  
CC The full length Factor VIII:C cDNA has two changes with respect to the  
CC published sequence (EPO application 160457):  
CC C to T at CNA at Leu 242 and T to C to C to C change at amino acid residue 1880  
CC (Phe to Leu). The product is produced in approx. 20 times higher  
CC yields than previous recombinant produced factor VIII:C and are more  
CC easily purified. The peptide is used for treating haemophilia A, both  
CC acute and prolonged bleeding.  
CC See also AAN80444 and AAN80447.  
XX  
SQ Sequence 4275 BP; 1245 A; 940 C; 946 G; 1144 T; 0 other;  
  
Query Match 87.5%; Score 4171.6; DB 9; Length 4275;  
Best Local Similarity 98.0%; Pred. No. 0;  
Matches 4268; Conservative 0; Mismatches 4; Indels 84; Gaps 1;  
  
QY 328 gccacccaagataactactcctgtgctgcagtggaactgcatgtgactatgcaaatgcat 387  
|||||  
Db 4 gccacccaagataactactcctgtgctgcagtggaactgcatgtgactatgcaaatgcat 63  
QY 388 ctggtgagctgctgtgacgcaagaattctccttagagtgccaatctttccattc 447  
|||||  
Db 64 ctggtgagctgctgtgacgcaagaattctccttagagtgccaatctttccattc 123  
QY 448 aacacccatgctgtatcaaaaagactctgtttgaatcaatcaggttcaacctttcaac 507  
|||||  
Db 124 aacacccatgctgtatcaaaaagactctgtttgaatcaatcaggttcaacctttcaac 183  
QY 508 atcgtaagccaagggcacccttgatgtgtgtgtagtcttaccatcccaagctgaggtt 567  
|||||  
Db 184 atcgtaagccaagggcacccttgatgtgtgtgtagtcttaccatcccaagctgaggtt 243  
QY 568 tatgatacagtggtcattacacttaagaacatggtcttccatctgtcagttcattagct 627  
|||||  
Db 244 tatgatacagtggtcattacacttaagaacatggtcttccatctgtcagttcattagct 303  
QY 628 gttgtgtatctctactctgaaagctctctgagggagctgaatatgtatgcatcagaccagtc 687  
|||||  
Db 304 gttgtgtatctctactctgaaagctctctgagggagctgaatatgtatgcatcagaccagtc 363  
QY 688 agggagaagaagaatgataaagcttccctgtgtggaagccaataatgtctgtgcagctc 747  
|||||  
Db 364 agggagaagaagaatgataaagcttccctgtgtggaagccaataatgtctgtgcagctc 423  
QY 748 ctgaaagaagaatggttccaatggtccttgaccactgtgcttacttactatcttct 807  
|||||  
Db 424 ctgaaagaagaatggttccaatggtccttgaccactgtgcttacttactatcttct 483  
QY 808 catgtgacctggttaaaaagcttgaattcagggcttatttggagccctactagatgtaga 867  
|||||  
Db 484 catgtgacctggttaaaaagcttgaattcagggcttatttggagccctactagatgtaga 543  
QY 868 gaaaggagctgtgccaagaaagacaagacagccttgcaacaatttactacttctgct 927  
|||||  
Db 544 gaaaggagctgtgccaagaaagacaagacagccttgcaacaatttactacttctgct 603  
QY 928 gtattgtgaaaggaaagtgtgacctcagaacaaagaatcctgtatgagatagag 987  
|||||  
Db 604 gtattgtgaaaggaaagtgtgacctcagaacaaagaatcctgtatgagatagag 663  
QY 988 gatgtgcattctgtctggtcgtgcttcaaaaatgacacagtcgaatgttattgtataaagg 1047  
|||||  
Db 664 gatgtgcattctgtctggtcgtgcttcaaaaatgacacagtcgaatgttattgtataaagg 723  
QY 1048 tctctgcagctgcatgtgattgagtcacaggaataacagtcattatgagctgtgagatgag 1107  
|||||  
Db 724 tctctgcagctgcatgtgattgagtcacaggaataacagtcattatgagctgtgagatgag 783  
QY 1108 ggcacacactcctgaagtgcatcaataattcctcgaaggtacacacattctgttgaagaa 1167  
|||||  
Db 784 ggcacacactcctgaagtgcatcaataattcctcgaaggtacacacattctgttgaagaa 843



QY 1168 catgcgcagcgctcccttggnaaatctcgcaataacttctctactgtctcaaaactcttg 1227  
|||||  
Db 844 catgcgcagcgctcccttggnaaatctcgcaataacttctctactgtctcaaaactcttg 903  
QY 1228 atggaaccttggacagttctctactgttttgcatactcttcccaacaatgagtgcag 1287  
|||||  
Db 904 atggaaccttggacagttctctactgttttgcatactcttcccaacaatgagtgcag 963  
QY 1288 gaagcttatgtcagaagttagacagctgtctcagaaggaaaccccaactcgaaatgnaaataat 1347  
|||||  
Db 964 gaagcttatgtcagaagttagacagctgtctcagaaggaaaccccaactcgaaatgnaaataat 1023  
QY 1348 gaaagaacggaagagactatgatgatcttactgtattctgnaattgagtggtcgagttt 1407  
|||||  
Db 1024 gaagaagcggaagagactatgatgatcttactgtattctgnaattgagtggtcgagttt 1083  
QY 1408 gatgatgaacaactctctctctcttcccaaatctgctcaagtctgcagaagaagcatccctaa 1467  
|||||  
Db 1084 gatgatgaacaactctctctctcttcccaaatctgctcaagtctgcagaagaagcatccctaa 1143  
QY 1468 actgtgggtacatctacatgtctgtctgaagaaggagagctggaactatgtcccttaactctc 1527  
|||||  
Db 1144 actgtgggtacatctacatgtctgtctgaagaaggagagctggaactatgtcccttaactctc 1203  
QY 1528 gcccccgaatgacagaagttatataaagttcaataattgnaacaatggtccctcagcgagttggt 1587  
|||||  
Db 1204 gcccccgaatgacagaagttatataaagttcaataattgnaacaatggtccctcagcgagttggt 1263  
QY 1588 aggaagttacaaaaaagttccgattttatggtacatacacagatgnaaacttgaagctcgtgaa 1647  
|||||  
Db 1264 aggaagttacaaaaaagttccgattttatggtacatacacagatgnaaacttgaagctcgtgaa 1323  
QY 1648 gctatcgaagctatgaatcagaagatcttgggaacttacttatgtgggaagtgtgagacaca 1707  
|||||  
Db 1324 gctatcgaagctatgaatcagaagatcttgggaacttacttatgtgggaagtgtgagacaca 1383  
QY 1708 cgtgtgatattatatttaagaatcaagaacgagacacataaatactacccctcaggaatc 1767  
|||||  
Db 1384 cgtgtgatattatatttaagaatcaagaacgagacacataaatactacccctcaggaatc 1443  
QY 1768 actgatgtccgtctctctgttatctcaagaagattaccaaaagtggttaaaacatttgaagat 1827  
|||||  
Db 1444 actgatgtccgtctctctgttatctcaagaagattaccaaaagtggttaaaacatttgaagat 1503  
QY 1828 ttcccaattcgcagaggaataattcaaatataatggaacgtagacgttagaagatgg 1887  
|||||  
Db 1504 ttcccaattcgcagaggaataattcaaatataatggaacgtagacgttagaagatgg 1563  
QY 1888 ccaactaatatcagatcctcgtgtgcctgcagccgctatlaactctagtttcgttaataatggag 1947  
|||||  
Db 1564 ccaactaatatcagatcctcgtgtgcctgcagccgctatlaactctagtttcgttaataatggag 1623  
QY 1948 agagatcttagcttcaagagatctattgtgcccctctccctcactctgtctcaaaaagaactgtat 2007  
|||||  
Db 1624 agagatcttagcttcaagagatctattgtgcccctctccctcactctgtctcaaaaagaactgtat 1683  
QY 2008 caaagaagaaacagataatgttcasagaacaggaatgtcaltccctgttttctgtattgtat 2067  
|||||  
Db 1684 caaagaagaaacagataatgttcasagaacaggaatgtcaltccctgttttctgtattgtat 1743  
QY 2068 gagaacccaagctgtgtactcctacagagaataatacaacgcttcttcccaatccagctgga 2127  
|||||  
Db 1744 gagaacccaagctgtgtactcctacagagaataatacaacgcttcttcccaatccagctgga 1803  
QY 2128 gtgcagcttgaagatccagagattccaaagctccaaacatcagacagacataatgtgctat 2187  
|||||  
Db 1804 gtgcagcttgaagatccagagattccaaagctccaaacatcagacagacataatgtgctat 1863  
QY 2188 gtctttgatagtttcaggttgcagttgtttgtgcataaggtgcagctactgtgtacattta 2247  
|||||  
Db 1864 gtctttgatagtttcaggttgcagttgtttgtgcataaggtgcagctactgtgtacattta 1923

QY 2248 agcaattggagcacagactgacttctctctctgtctctctctctctctctctctcaaacac 2307  
|||||  
Db 1924 agcaattggagcacagactgacttctctctctgtctctctctctctctctctctcaaacac 1983  
QY 2308 aaaatggtctatgaaagacaactcaactctctctctctctctctctctctctctctctcaaac 2367  
|||||  
Db 1984 aaaatggtctatgaaagacaactcaactctctctctctctctctctctctctctctctcaaac 2043  
QY 2368 tcgattggaaaaacccaggtctatggaattctgtgggtgcacaaactcaagacttctcgaaacaga 2427  
|||||  
Db 2044 tcgattggaaaaacccaggtctatggaattctgtgggtgcacaaactcaagacttctcgaaacaga 2103  
QY 2428 ggcacatgcgcctctactgaaagtttctagtgtgacaaagaacacgtgtatlatcagag 2487  
|||||  
Db 2104 ggcacatgcgcctctactgaaagtttctagtgtgacaaagaacacgtgtatlatcagag 2163  
QY 2488 gacaaattatgaagatatttcagcatactgtctgtgataaaaaaatgcatatgaaaccaa 2547  
|||||  
Db 2164 gacaaattatgaagatatttcagcatactgtctgtgataaaaaaatgcatatgaaaccaa 2223  
QY 2548 agctctccagaattcgaagacaccttagcaactagcgaaagaactttaatgcccacca 2607  
|||||  
Db 2224 ----- 2223  
QY 2608 ccagctcttgaaacgcatcaacgaggaataactcgtactactctcagtcagatcaagaag 2667  
|||||  
Db 2224 ----- gaataactcgtactactctcagtcagatcaagaag 2259  
QY 2668 gaatttgactatgatgatataccatcatcagttgaatgaaagaaagaaatttgaacattat 2727  
|||||  
Db 2260 gaatttgactatgatgatataccatcatcagttgaatgaaagaaagaaatttgaacattat 2319  
QY 2728 gctgaggtatgaaatcagagcccccagcttccaaagaacaaacagacactattat 2787  
|||||  
Db 2320 gctgaggtatgaaatcagagcccccagcttccaaagaacaaacagacactattat 2379  
QY 2788 gctgaggtatgaaaggtcctgtgattatggaatgagtagtcccaatggttctaagaac 2847  
|||||  
Db 2380 gctgaggtatgaaaggtcctgtgattatggaatgagtagtcccaatggttctaagaac 2439  
QY 2848 aggtctcagaggtgaggtgctctcctcagttcaagaagttgtttccaggaattactgat 2907  
|||||  
Db 2440 aggtctcagaggtgaggtgctctcctcagttcaagaagttgtttccaggaattactgat 2499  
QY 2908 gctcctcttactcagcccttataccgtgagagaactaaattggaacatttggagctcctggg 2967  
|||||  
Db 2500 gctcctcttactcagcccttataccgtgagagaactaaattggaacatttggagctcctggg 2559  
QY 2968 ccatataagaagcagaagttggaagataataatcatggttaacttccagaaatcagcctct 3027  
|||||  
Db 2560 ccatataagaagcagaagttggaagataataatcatggttaacttccagaaatcagcctct 2619  
QY 3028 cgtccctactctctctatcttagccttatttcttatgaggaagttcagaagcagaagca 3087  
|||||  
Db 2620 cgtccctactctctctatcttagccttatttcttatgaggaagttcagaagcagaagca 2679  
QY 3088 gaacctagaaaaaacttgttcaagcctatgaaacccaacttacttcttggaaagtgcga 3147  
|||||  
Db 2680 gaacctagaaaaaacttgttcaagcctatgaaacccaacttacttcttggaaagtgcga 2739  
QY 3148 catcatatggaacccaactaaagatgagttgactgtcgaagcctgtggtctatctctgat 3207  
|||||  
Db 2740 catcatatggaacccaactaaagatgagttgactgtcgaagcctgtggtctatctctgat 2799  
QY 3208 gttgacctggaaaaagatggtgactcagcagcctgtatttggacccttctgtgtcgcaact 3267  
|||||  
Db 2800 gttgacctggaaaaagatggtgactcagcagcctgtatttggacccttctgtgtcgcaact 2859  
QY 3268 aacacatgaaacccctgtcattcggagacaagtgtgacgtatgaggaatttgcctgttttc 3327  
|||||  
Db 2860 aacacatgaaacccctgtcattcggagacaagtgtgacgtatgaggaatttgcctgttttc 2919  
QY 3328 accatctttagatgagaccaaagctgttacttcaactgaaaaatagaaaaaactgcag 3387



```

Db 2920 accactttgttgagaccaaagcgtgacttcaactgtaaaatcgtgaagaagcagcgg 2979
|
QY 3388 gctccctgcaatatccagatggaagatcccacttttaaagaattatcgcttccatgca 3447
|
Db 2980 gctccctgcaatatccagatggaagatcccacttttaaagaattatcgcttccatgca 3039
|
QY 3448 atcaatgctcacaatatgatacactactgcttagtaatgctcagaatcagaagatt 3507
|
Db 3040 atcaatgctcacaatatgatacactactgcttagtaatgctcagaatcagaagatt 3099
|
QY 3508 cgaatgctatcgtctcagcatggtgagcaatgaaacatcattatcatttcaatgga 3567
|
Db 3100 cgaatgctatcgtctcagcatggtgagcaatgaaacatcattatcatttcaatgga 3159
|
QY 3568 catggttccatctgacgaaaaaagagagatataaatggaactgtacactctatcca 3627
|
Db 3160 catggttccatctgacgaaaaaagagagatataaatggaactgtacactctatcca 3219
|
QY 3628 ggtgtttttgagacagtggaatgttaacatccaaagctggaatttgccggttggaatgc 3687
|
Db 3220 ggtgtttttgagacagtggaatgttaacatccaaagctggaatttgccggttggaatgc 3279
|
QY 3688 ctatgtgagacatctacatgctgagatgagcaacttttctggtgatacagaataag 3747
|
Db 3280 ctatgtgagacatctacatgctgagatgagcaacttttctggtgatacagaataag 3339
|
QY 3748 tgtcagactccctcggggaatggtctctggaacattgagaatttcagatacagctca 3807
|
Db 3340 tgtcagactccctcggggaatggtctctggaacattgagaatttcagatacagctca 3399
|
QY 3808 ggaacataatgacagtgaggcccaagctggccagacttcaattccggatcacaat 3867
|
Db 3400 ggaacataatgacagtgaggcccaagctggccagacttcaattccggatcacaat 3459
|
QY 3868 gctcgtgagaccaaaggagcccttcttggatcaaggctgacgtctggtgaccaaagatt 3927
|
Db 3460 gctcgtgagaccaaaggagcccttcttggatcaaggctgacgtctggtgaccaaagatt 3519
|
QY 3928 attcagagcatcaagaccaggggtgcccggtcagaagtcttccaaacctatactctcag 3987
|
Db 3520 attcagagcatcaagaccaggggtgcccggtcagaagtcttccaaacctatactctcag 3579
|
QY 3988 ttatcatcatgatagttctgtgatggaagaagtgagcagacttactgaggaattccact 4047
|
Db 3580 ttatcatcatgatagttctgtgatggaagaagtgagcagacttactgaggaattccact 3639
|
QY 4048 ggaaccttaatggtctctcttggcaatggtgattcatctggataaaacacataatttt 4107
|
Db 3640 ggaaccttaatggtctctcttggcaatggtgattcatctggataaaacacataatttt 3699
|
QY 4108 aacctccaattatgctcgatacatccgtttgacaccaactcaattatagatccgagc 4167
|
Db 3700 aacctccaattatgctcgatacatccgtttgacaccaactcaattatagatccgagc 3759
|
QY 4168 accttcgacatgaggtgattggtggtgatttaattggttcacatgacatggagattg 4227
|
Db 3760 accttcgacatgaggtgattggtggtgatttaattggttcacatgacatggagattg 3819
|
QY 4228 gagaagtaagaatatacatagatgacagatctactgcttccatcctcttaccataatgtt 4287
|
Db 3820 gagaagtaagaatatacatagatgacagatctactgcttccatcctcttaccataatgtt 3879
|
QY 4288 ggcacccgtgtctctcttcaaaaagctcgacttcaacctcaaggaggaatgactgtgaga 4347
|
Db 3880 ggcacccgtgtctctcttcaaaaagctcgacttcaacctcaaggaggaatgactgtgaga 3939
|
QY 4348 cctcaggtgaaataatccaaaagatggtcgtgcaagtggacttccagaagacaataaagctc 4407
|
Db 3940 cctcaggtgaaataatccaaaagatggtcgtgcaagtggacttccagaagacaataaagctc 3999
|
QY 4408 acagagtaactactcagggagtaaaatctctgcttaccagcatgtatgtgaaagagcttc 4467
|

```

```

Db 4000 acagagtaactactcagggagtaaaatctctgcttaccagcatgtatgtgaagatlc 4059
|
QY 4468 ctcaatcccaagcagtcgaagatgagccatcaagtggactctcttttccgaatggcaagta 4527
|
Db 4060 ctcaatcccaagcagtcgaagatgagccatcaagtggactctcttttccgaatggcaagta 4119
|
QY 4528 aagtttttcagggaataatcaagactccttcacactggtggtgaactcctagaccaccg 4587
|
Db 4120 aagtttttcagggaataatcaagactccttcacactggtggtgaactccttagaccaccg 4179
|
QY 4588 ttactgactcgtaactccttgaattcaacccagagttggtgacacagatgacctgagg 4647
|
Db 4180 ttactgactcgtaactccttgaattcaacccagagttggtgacacagatgacctgagg 4239
|
QY 4648 atggaagttctggtgctcgagacagagactctac 4683
|
Db 4240 atggaagttctggtgctcgagacagagactctac 4275
|

```

## RESULT 13

AA080447  
ID AA080447 standard; DNA: 4272 BP.

AC AA080447;

DT 10-OCT-1990 (first entry)

DE Modified factor VIII:C sequence with the R740-D1658 deletion.

KM Modified factor VIII:C; haemophilia; procoagulant;

KM blood coagulation; RD deletion; ss.

OS Homo sapiens.

PN W08800831-A.

PD 11-FEB-1988.

PF 31-JUL-1987; 87MO-US01814.

PR 01-AUG-1986; 86US-0893375.

PA (BIOJ ) BIOGEN NV.

PI Pasek MP;

DR WPI; 1988-049866/07.

DR P-PSDB; AAP80268.

PT New DNA sequences encoding modified factor VIII:C - with deletion of DNA encoding maturation polypeptide, useful for high yield transformation.

PS Claim 3; Page 47-48-49-50; 97P; English.

CC The RD deletion removes the DNA from Ser 741 to Ser 1657. A major part of the sequence encoding the maturation polypeptide of

CC factor VIII:C is deleted, i.e. Gln 744 - Asp 1563.

CC The full length Factor VIII:C cDNA has two changes with respect to the published sequence (EPO application 160457):

CC C to G at Cys 242 and T to C at Cys 243 change at amino acid residue 1880 (the to Leu). The product is produced in approx. 20 times higher

CC yields than previous recombinant produced factor VIII:C and are more easily purified. The peptide is used for treating haemophilia A, both

CC acute and prolonged bleeding.

CC See also AA080444 and AA080446.

XX Sequence 4272 BP; 1243 A; 941 C; 944 G; 1144 T; 0 other;

Query Match 87.5%; Score 4170; DB 9; Length 4272;  
Best Local Similarity 98.0%; Pred. No. 0;  
Matches 4267; Conservative 0; Mismatches 5; Indels 84; Gaps 1;



QY	328	ggaacccgaagaatgatactctcctctggtggtgcagtggaacgctgatactggaatataatgaaagtat	387
Db	1	gccaaccgaagaatactatacctcgtggtgcagtggaactgtlcaatggtgacataatgaaagtat	60
QY	388	ctcggtgaagctgcctctggaagcgaagaatcttccctctagaagtgccaaaatctttccatctc	447
Db	61	ctcggtgaagctgcctctggaagcgaagaatcttccctctagaagtgccaaaatctttccatctc	120
QY	448	aacaaccctgaatcgltgtacaaaagaactctgtttgtlagaatccaaggttcaacctttcaac	507
Db	121	aacaacctgaatcgltgtacaaaagaactctgtttgtlagaatccaaggttcaacctttcaac	180
QY	508	atcgctgaagccaagggccaacctgtgaatggtctgtcagaagttccatacctcagctgaagtt	567
Db	181	atcgctgaagccaagggccaacctgtgaatggtctgtcagaagttccatacctcagctgaagtt	240
QY	568	tatgatatacagtggtgcatatacacttaagaagacatggcttccatctcctgtaagcttccatgct	627
Db	241	tatgatatacagtggtgcatatacacttaagaagacatggcttccatctcctgtaagcttccatgct	300
QY	628	gtctggtgtatcctactctggaagactcttgaggagctgaatatgatatgatacagaatcca	687
Db	301	gtctggtgtatcctactctggaagactcttgaggagctgaatatgatatgatacagaatcca	360
QY	688	agggaagagaagaagtgataaagttcttccctgtgtggaagccaataatgctgtcgaggtc	747
Db	361	agggaagagaagaagtgataaagttcttccctgtgtggaagccaataatgctgtcgaggtc	420
QY	748	ctgaagaagaatgataccaatgtgctcttgacccaactgtgctctactacatactatcttct	807
Db	421	ctgaagaagaatgataccaatgtgctcttgacccaactgtgctctactacatactatcttct	480
QY	808	catgttgacctgtgtaaaagacttgaatccaagctcataltgagccctactagtatgtaga	867
Db	481	catgttgacctgtgtaaaagacttgaatccaagctcataltgagccctactagtatgtaga	540
QY	868	gaaggagggtctctgcccgaaggaaaagacaagacgcttgccaacaattatatactatttctc	927
Db	541	gaaggagggtctctgcccgaaggaaaagacaagacgcttgccaacaattatatactatttctc	600
QY	928	gtatcttgaagggaaaaggtttggtcactcagaacaaaagaactcccttgatgcagataga	987
Db	601	gtatcttgaagggaaaaggtttggtcactcagaacaaaagaactcccttgatgcagataga	660
QY	988	gatagtctgacatctgctctggtggtccttgacctaaaatgtcaacagatgcaatgtgtatgtaaacgg	1047
Db	661	gatagtctgacatctgctctggtggtccttgacctaaaatgtcaacagatgcaatgtgtatgtaaacgg	720
QY	1048	tctctgtccaggtctgcatcttggaatgtgcacaggaatccaagctatcttgcatgtatgtgaatg	1107
Db	721	tctctgtccaggtctgcatcttggaatgtgcacaggaatccaagctatcttgcatgtatgtgaatg	780
QY	1108	ggcacaacctctggaagtgcatacaatalctctcgaaaggttcaacacatcttctgttgagagac	1167
Db	781	ggcacaacctctggaagtgcatacaatalctctcgaaaggttcaacacatcttctgttgagagac	840
QY	1168	catctgcagaagcgctccttgggaatcttcgccaataacttctcttactgtctcaaacactctg	1222
Db	841	catctgcagaagcgctccttgggaatcttcgccaataacttctcttactgtctcaaacactctg	900
QY	1228	atggaaccttggaagtttctacagtcttctgtcatatctcttcccaaccaatgatagtgacatg	1287
Db	901	atggaaccttggaagtttctacagtcttctgtcatatctcttcccaaccaatgatagtgacatg	960
QY	1288	gaagcctatgtlcaaaagtagaacagctgttccagaaggaaacccaactagaatgaaaaatat	1347
Db	961	gaagcctatgtlcaaaagtagaacagctgttccagaaggaaacccaactagaatgaaaaatat	1020
QY	1348	gaaggaagcggaaagactatgtatgataccttactcgtatcttgaaatggatgtgtcaggttt	1407
Db	1021	gaaggaagcggaaagactatgtatgataccttactcgtatcttgaaatggatgtgtcaggttt	1080
QY	1408	gatagtatacaactctctcttataccaatatcgctcagttgtgcagaagcatcctaata	1467

Db	1081	gatgatgcaactctctctctcttcaatcccaattcgcctcagttgccaagaagatccctaa	1140
Qy	1468	acttgggtacacttaacatgctgctgcaagagagagactggnactatgctctcccttaagctc	1527
Db	1141	acttgggtacacttaacatgctgctgcaagagagagactggnactatgctctcccttaagctc	1200
Qy	1528	gccccgactgcaagaagtattaaagtcacatttgacaatgugccctcagaggtatgct	1587
Db	1201	gccccgactgcaagaagtattaaagtcacatttgacaatgugccctcagaggtatgct	1260
Qy	1588	aggaagtacaaaagaatcccgattatgagcaatgacagatgaaaccccttaagctctgaa	1647
Db	1261	aggaagtacaaaagaatcccgattatgagcaatgacagatgaaaccccttaagctctgaa	1320
Qy	1648	gctattcagcaatgaatcagaagatcttggagcccttaacttattatgugggaaagtgtgagacaca	1707
Db	1321	gctattcagcaatgaatcagaagatcttggagcccttaacttattatgugggaaagtgtgagacaca	1380
Qy	1708	ctgtgtgattatattaaagaatcgaagcagaagcagacatatcaacttaccctcagaaatc	1767
Db	1381	ctgtgtgattatattaaagaatcgaagcagaagcagacatatcaacttaccctcagaaatc	1440
Qy	1768	actgaatgcgtctcccttggatttcaagaagagattacaaaagtggtataaacatttgaaat	1827
Db	1441	actgaatgcgtctcccttggatttcaagaagagattacaaaagtggtataaacatttgaaat	1500
Qy	1828	tttccaattctgccaagagagaataattcaaatataaatggacagctgactgtgagaagaatgg	1887
Db	1501	tttccaattctgccaagagagaataattcaaatataaatggacagctgactgtgagaagaatgg	1560
Qy	1888	ccaactaatatgaatctctctggtgctgcgaagccgctatctactcagtttctgttaatatgag	1947
Db	1561	ccaactaatatgaatctctctggtgctgcgaagccgctatctactcagtttctgttaatatgag	1620
Qy	1948	agaagctcagcttcaaggaacatcttggagccctccctatcctgctacataaaaagactgtagat	2007
Db	1621	agaagctcagcttcaaggaacatcttggagccctccctatcctgctacataaaaagactgtagat	1680
Qy	2008	caaaagaggaacccagataatgtctcagacagaaggaatgtcaatccctgtttctgttattgat	2067
Db	1681	caaaagaggaacccagataatgtctcagacagaaggaatgtcaatccctgtttctgttattgat	1740
Qy	2068	gagaacccaagaactgtgataccctcagaagaaatatacaacgctttctcccaatccagacttga	2127
Db	1741	gagaacccaagaactgtgataccctcagaagaaatatacaacgctttctcccaatccagacttga	1800
Qy	2128	gtgcagactgtgagatccagagatcccaagcctccaacataatgacaagactcaaatgtgcata	2187
Db	1801	gtgcagactgtgagatccagagatcccaagcctccaacataatgacaagactcaaatgtgcata	1860
Qy	2188	gtttttgatagtttgtcagactgtgcagattgtttgtgcatgagtgagcatactgttatactta	2247
Db	1861	gtttttgatagtttgtcagactgtgcagattgtttgtgcatgagtgagcatactgttatactta	1920
Qy	2248	agcatttggagaagcagactgactctctcttctgctctcttctctcgtgatacttccaacac	2307
Db	1921	agcatttggagaagcagactgactctctcttctgctctcttctctcgtgatacttccaacac	1980
Qy	2308	aaaatgtgctctatgagaacacacatcccttatcccatctctcagaggaagaactgtcttcatg	2367
Db	1981	aaaatgtgctctatgagaacacacatcccttatcccatctctcagaggaagaactgtcttcatg	2040
Qy	2368	tcgatgtgaaaacccagatctatgagatcttgygggtgccaacactcagacttctggaacaga	2427
Db	2041	tcgatgtgaaaacccagatctatgagatcttgygggtgccaacactcagacttctggaacaga	2100
Qy	2428	ggcagatcccgctcttaccgaaggtttctagtgttgacaagaagaacggtgattattacagag	2487
Db	2101	ggcagatcccgctcttaccgaaggtttctagtgttgacaagaagaacggtgattattacagag	2160
Qy	2488	gacagttatgagaatatttcaagatacttctgagcaaaaacatgccaattgaaccaga	2547



Db 2161 gacagttatgaagatalatttcagcalacttgcgtgagtaaaaaacatgccaattgacaacaaga 2220  
QY 2548 agctctccagaattcaagaacacctagcacatagcaaaaaagcaatttaattgcacacca 2607  
Db 2221 ----- 2220  
QY 2608 ccaagtctgaacgcacacaacggaataactgcctaactctcttcagtcagatccaagag 2667  
Db 2221 ----- gaaataaccgtactactctcttcagtcagatccaagag 2256  
QY 2668 gaattgactatgatagtacacataacagttgaaatgagaagaagaatttgacattat 2127  
Db 2257 gaattgactatgatagtacacataacagttgaaatgagaagaagaatttgacattat 2316  
QY 2728 gatgagatgaataatcagagccccgcagcttccaagaagaacacagaccctttat 2787  
Db 2317 gatgagatgaataatcagagccccgcagcttccaagaagaacacagaccctttat 2376  
QY 2788 gctgcagtgagagagctcttggaattatggaatgagtagcctccacacatgcttaagaac 2847  
Db 2377 gctgcagtgagagagctcttggaattatggaatgagtagcctccacacatgcttaagaac 2436  
QY 2848 agggctcagaagtgcagtgctccctcagttcagaagaagtgtttccaggaattactgat 2907  
Db 2437 agggctcagaagtgcagtgctccctcagttcagaagaagtgtttccaggaattactgat 2496  
QY 2908 ggtcccttctcagccttatccgttgcgtaagaactaattgaaactttggaccctcgggg 2967  
Db 2497 ggtcccttctcagccttatccgttgcgtaagaactaattgaaactttggaccctcgggg 2556  
QY 2968 ccataataagacagagaagtctgaagataatcatgtaacttccagaatccagcctc 3027  
Db 2557 ccataataagacagagaagtctgaagataatcatgtaacttccagaatccagcctc 2616  
QY 3028 cgtccctatctcttattcttaagccttattcttataagagaatcaagagcaaggagca 3087  
Db 2617 cgtccctatctcttattcttaagccttattcttataagagaatcaagagcaaggagca 2676  
QY 3088 gaacctagaaaaaactttgtcagagcctaagtgaacccaataacttttggaaagtgcac 3147  
Db 2677 gaacctagaaaaaactttgtcagagcctaagtgaacccaataacttttggaaagtgcac 2736  
QY 3148 catcataatgacacccaactaaagatgagtttgaactgcacaaacccctgggcttattcctgat 3207  
Db 2737 catcataatgacacccaactaaagatgagtttgaactgcacaaacccctgggcttattcctgat 2796  
QY 3208 gtgcacctgaaaaaagatgtgcactcagcctgattgagaccctctcgtctgcacact 3267  
Db 2797 gtgcacctgaaaaaagatgtgcactcagcctgattgagaccctctcgtctgcacact 2856  
QY 3268 aacacactgaaacccctgtcactgagagaacagtgaacagtaagaatttgcctgcttttc 3327  
Db 2857 aacacactgaaacccctgtcactgagagaacagtgaacagtaagaatttgcctgcttttc 2916  
QY 3328 accatcttcatagagccaacaaagctgttacttcaactgaaataatggaaaaaacctcag 3387  
Db 2917 accatcttcatagagccaacaaagctgttacttcaactgaaataatggaaaaaacctcag 2976  
QY 3388 gctccctgcaataatccagatgagaatccacatttaagaagaatatcgcctcactga 3447  
Db 2977 gctccctgcaataatccagatgagaatccacatttaagaagaatatcgcctcactga 3036  
QY 3448 ataaatgagcaataatgatactacacacgttgtagtagtgcagtaagaataaagatt 3507  
Db 3037 ataaatgagcaataatgatactacacacgttgtagtagtgcagtaagaataaagatt 3096  
QY 3508 cgaatgtatctgcagcatgaggcagcaaatgaaacaatccatcttacttcaatgag 3567  
Db 3097 cgaatgtatctgcagcatgaggcagcaaatgaaacaatccatcttacttcaatgag 3156  
QY 3568 catgtgttcaactgttacgaaaaaagaggtataaaatgagcactgtacaatctctac 3627  
Db 3157 catgtgttcaactgttacgaaaaaagaggtataaaatgagcactgtacaatctctac 3216

QY 3628 ggtcttttggagacagtggaatgtttacacatccaaaagctggaatttggcgggtggaatgc 3687  
Db 3217 ggtcttttggagacagtggaatgtttacacatccaaaagctggaatttggcgggtggaatgc 3276  
QY 3688 ctatattgagagacatctacatgcttgagatgagcacacttttctgtgtgtaacaagaatga 3747  
Db 3277 ctatattgagagacatctacatgcttgagatgagcacacttttctgtgtgtaacaagaatga 3336  
QY 3748 tctcagaatcccttggaatgagcttcttgacacattagagaatttcaagatccagcttca 3807  
Db 3337 tctcagaatcccttggaatgagcttcttgacacattagagaatttcaagatccagcttca 3396  
QY 3808 ggaacaatgagagagtgggcccaaaagcttgcaggaattatattccggaataatcaat 3867  
Db 3397 ggaacaatgagagagtgggcccaaaagcttgcaggaattatattccggaataatcaat 3456  
QY 3868 gctctgagacacaaagagacccttctcttgatcaagtgatctgttgcaccaatgat 3927  
Db 3457 gctctgagacacaaagagacccttctcttgatcaagtgatctgttgcaccaatgat 3516  
QY 3928 attcaagcatcaagaacagaggtgcctgcagaaagtctccagcctctacatctcag 3987  
Db 3517 attcaagcatcaagaacagaggtgcctgcagaaagtctccagcctctacatctcag 3576  
QY 3988 ttatatacatgfatagcttcttgatggagaagaatgagagacattatcgagaattccact 4047  
Db 3577 ttatatacatgfatagcttcttgatggagaagaatgagagacattatcgagaattccact 3636  
QY 4048 ggaaccttaatggtctctcttcttgacaatgttgatcatcatctggaataaaacaatatctt 4107  
Db 3637 ggaaccttaatggtctctcttcttgacaatgttgatcatcatctggaataaaacaatatctt 3696  
QY 4108 aacctccaattatgtctcogatacatccgttgcacccaactccattatagcatctcagc 4167  
Db 3697 aacctccaattatgtctcogatacatccgttgcacccaactccattatagcatctcagc 3756  
QY 4168 actcttcgcagtggaagtgttgagctgttgatgaatagttgagcagcatctgaggaatg 4227  
Db 3757 actcttcgcagtggaagtgttgagctgttgatgaatagttgagcagcatctgaggaatg 3816  
QY 4228 gagaagtaaaagcaatcatcagaatgacagatatactgcttccacttcaacttaacaaatgatt 4287  
Db 3817 gagaagtaaaagcaatcatcagaatgacagatatactgcttccacttcaacttaacaaatgatt 3876  
QY 4288 gccacctgtctctcttccaaaagctgcactacactccaaggagaggttaatgcctggaga 4347  
Db 3877 gccacctgtctctcttccaaaagctgcactacactccaaggagaggttaatgcctggaga 3936  
QY 4348 cctcaggtgaataatccaagaagtggttgcagtggaacttccaagaagaatgaaatgc 4407  
Db 3937 cctcaggtgaataatccaagaagtggttgcagtggaacttccaagaagaatgaaatgc 3996  
QY 4408 acagagtaactactcaagggagtaaaatcctctgcttaccagcatgtatgtgaagagttc 4467  
Db 3997 acagagtaactactcaagggagtaaaatcctctgcttaccagcatgtatgtgaagagttc 4056  
QY 4468 ctcaatctcaagcagtcagaatgagcatcagtggaactctcttcttcaagaatggcaaatga 4527  
Db 4057 ctcaatctcaagcagtcagaatgagcatcagtggaactctcttcttcaagaatggcaaatga 4116  
QY 4528 aaggttttccagggaaataaagaactcttccacacccgttgtaactcttagaccacag 4587  
Db 4117 aaggttttccagggaaataaagaactcttccacacccgttgtaactcttagaccacag 4176  
QY 4588 ttactgactcgcactactcgaattccaccccaagagttggttgacacagaattgcctcagag 4647  
Db 4177 ttactgactcgcactactcgaattccaccccaagagttggttgacacagaattgcctcagag 4236  
QY 4648 atggaggttctggtctgcagagcaggaactctac 4683  
Db 4237 atggaggttctggtctgcagagcaggaactctac 4272



RESULT 14

AAx82258  
ID AAX82258 standard; cDNA: 4373 BP.

XX AAX82258;

XX 18-AUG-1999 (first entry)

DE Beta-domain deleted Factor VIII protein encoding cDNA.

XX Factor VIII protein; gene modification; gene therapy; clinical disorder;

KW splicing pattern; RNA processing; gene regulation; beta-domain; human;

XX ss.

OS Homo sapiens.

PN W0929848-A1.

PD 17-JUN-1999.

PF 25-NOV-1998: 98MO-US25354.

PR 16-JAN-1998: 98US-0071596.

PR 05-DEC-1997: 97US-0067614.

PA (IMMU-) IMMUNE RESPONSE CORP.

PI Bidlingmaier S, Gonzales JEN, Ill CR, Yang CQ:

XX WPI: 1999-385602/32.

DR P-PSDB; AAY21675.

PT Genes and vectors exhibiting increased expression and novel splicing

PT patterns, useful for expression of, e.g. beta-domain deleted factor

PT VIII

XX Claim 9: Page 72-78; 123pp; English.

CC The invention describes novel genes and vectors exhibiting increased

CC expression and novel splicing patterns. It provides a gene encoding a

CC Factor VIII protein, that comprises one or more consensus or near

CC consensus splice sites which have been corrected to increase expression.

CC The method, DNA sequences and expression vectors can be used to increase

CC the expression of a gene, especially a Factor VIII gene. Genes containing

CC modified 5' and/or 3' untranslated regions have optimized expression

CC levels and tissue-specific expression. The methods are used for

CC identification and correction of consensus splice sites, addition of

CC introns, optimization of 5' and 3' untranslated regions and increase in

CC cytoplasmic RNA accumulation. Hence the DNAs are useful in gene therapy

CC to treat a clinical disorder, to study RNA processing and/or gene

CC regulation. The present sequence represents a cDNA encoding a beta-domain

CC deleted factor VIII protein.

XX Sequence 4373 BP; 1278 A; 1011 C; 938 G; 1146 T; 0 other;

Query Match 87.3%; Score 4161; DB 20; Length 4373;

Best Local Similarity 96.8%; Pred. No. 0;

Matches 42/3; Conservative 0; Mismatches 100; Indels 42; Gaps 1;

QY 271 atgcaaatagagctctcaccctgtcttctgtgcttctgtgcatctgtcttagtgc 330

DB 1 atggaatatagagctctcaccctgtcttctgtgcttctgtgcatctgtcttagtgc 60

QY 331 accagaagatactactcctgggtgcagtggaactgtcatgagagataatgcaagtatc 390

DB 61 accagaagatactactcctgggtgcagtggaactgtcatgagagataatgcaagtatc 120

QY 391 ggtgagctgctgtgagcgaagattcctcctcctagtagtgcacaaatcttccattcac 450

DB 121 ggtgagctgctgtgagcgaagattcctcctcctcctgctgctgcaaatcttccattcac 180

QY 451 acctcagctcgtgtacaaaaaagactctgtttgttagaatccaggttccacctttcaacatc 510

DB 181 acctcagctcgtgtacaaaaaagactctgtttgttagaatccaggttccacctttcaacatc 240

QY 511 gctaaagcgaagccacccctgtgattgtctgtatgcttcacatccagctgaggtttat 570

DB 241 gctaaagcgaagccacccctgtgattgtctgtatgcttcacatccagctgaggtttat 300

QY 571 gatacagtggtcattacacttaagaacatgtctcccatcttcctcagcttcatgtctgt 630

DB 301 gatacagtggtcattacacttaagaacatgtctcccatcttcctcagcttcatgtctgt 360

QY 631 ggttalcctcacttgtaaaagctctcgtgagagcttgataatgatgatcagacagtc 690

DB 361 ggttalcctcacttgtaaaagctctcgtgagagcttgataatgatgatcagacagtc 420

QY 691 gagaagaagaatgataagctcttccctcgtgtggaagccatcatatgtctgtgcagctc 750

DB 421 gagaagaagaatgataagctcttccctcgtgtggaagccatcatatgtctgtgcagctc 480

QY 751 aaagagaatgtgccaatgagcctcagccactgtgaccttgacttactatcttctcat 810

DB 481 aaagagaatgtgccaatgagcctcagccactgtgaccttgacttactatcttctcat 540

QY 811 gtagacctgtgtataaagacttgaaatccagccctcatgtgagccctactagatagagaa 870

DB 541 gtagacctgtgtataaagacttgaaatccagccctcatgtgagccctactagatagagaa 600

QY 871 ggtgagctgtgcccagaagaaagacacagacttgacaaatattactactcttctgtc 930

DB 601 ggtgagctgtgcccagaagaaagacacagacttgacaaatattactactcttctgtc 660

QY 931 ttgatgaaggaaagttgtgcatctcagaagaacagaccccttgatcagatagagat 990

DB 661 ttgatgaaggaaagttgtgcatctcagaagaacagaccccttgatcagatagagat 720

QY 991 gctgcatctgtctcggccttggtcctaaatgtcacacagctgattatgtataaagagct 1050

DB 721 gctgcatctgtctcggccttggtcctaaatgtcacacagctgattatgtataaagagct 780

QY 1051 ctgcgaagctctgattgtagatccacagaagaatcagctcattgtgcattgtatgtaagggc 1110

DB 781 ctgcgaagctctgattgtagatccacagaagaatcagctcattgtgcattgtatgtaagggc 840

QY 1111 acctccttgaaagcactcaatatctcctcgaaggtccacacattctcttggaagaacct 1170

DB 841 acctccttgaaagcactcaatatctcctcgaaggtccacacattctcttggaagaacct 900

QY 1171 cgcagagctctcttggaatctgcacaaataacttctctactcgtctcaaacactctgatg 1230

DB 901 cgcagagctctcttggaatctgcacaaataacttctctactcgtctcaaacactctgatg 960

QY 1231 gaccttgacaagttcttctacgtttttgtcatatctctcccaacaagtgtgcagtgaa 1290

DB 961 gaccttgacaagttcttctacgtttttgtcatatctctcccaacaagtgtgcagtgaa 1020

QY 1291 gcttatgtcacaagtagacagctgtccagaggaaccccaactacgaatgtaaaaaatgtaa 1350

DB 1021 gcttatgtcacaagtagacagctgtccagaggaaccccaactacgaatgtaaaaaatgtaa 1080

QY 1351 gaagcggaaagactatgatatgatacttactgattctgaaatgtagtgtgcaggtttgat 1410

DB 1081 gaagcggaaagactatgatatgatacttactgattctgaaatgtagtgtgtcagattgat 1140

QY 1411 gatgaacaactctccttcttattccaaatctcgctcagttgtccaaagaagatccttaaac 1470

DB 1141 gatgaacaactctccttcttattccaaatctcgctcagttgtccaaagaagatccttaaac 1200

QY 1471 tgggtacattacattgtctgcaggaagagagactcgggaagactctcccttagctccggcc 1530

DB 1201 tgggtacattacattgtctgcaggaagagagactcgggaagactctcccttagctccggcc 1260

QY 1531 ccgcatgacagaagttataaaagtcaatatttgaacaatgtgcctcagcgagattgttagg 1590



|||||  
Db 1261 ccgagatcaggaagttataaagccaatttgacaatgcccctcagcgattggaag 1320  
QY 1591 aagracaaaaaagtcgagattatggtacacagatgaaacctttaagactcgtgaagct 1650  
Db 1321 aagtaacaaaaagtcgagattatggtacacagatgaaacctttaagactcgtgaagct 1380  
QY 1651 attagatgaatcagaagatccttgagaccttctaattatgagggaagtgtgagacacgt 1710  
Db 1381 attcagatgaatcagaagatccttgagaccttctaattatgagggaagtgtgagacacgt 1440  
QY 1711 ttgattatattaagaatcagaagacacatataacatctacccctcaggaatcact 1770  
Db 1441 ctcatattattaagaatcagaagacacatataacatctacccctcaggaatcacc 1500  
QY 1771 gatgctcgctcctttgttatccaaggagattaccaaaggttaaaacatttgaagattt 1830  
Db 1501 gatgctcgctcctttgttatccaaggagattaccaaaggttaaaacatttgaagattt 1560  
QY 1831 ccaatttcgacgagaagaatattccaataataaatgacagtgactgtagaagatgagcca 1890  
Db 1561 ccaatttcgacgagaagaatattccaataataaatgacagtgactgtagaagatgagcca 1620  
QY 1891 actaaatcagatcctcgttgctgacccgctatctactagtttcgttaatatgagaga 1950  
Db 1621 actaaatcagatcctcgttgctgacccgctatctactagtttcgttaatatgagaga 1680  
QY 1951 gatctagctcagagactcaatctggtcctctccatcctgctacaaagaatctgtatgacaa 2010  
Db 1681 gatctagctcagagactcaatctggtcctctccatcctgctacaaagaatctgtatgacaa 1740  
QY 2011 aggggaacacagataatgtcagacaagaggaaatgctactcgtttcttgatattgagag 2070  
Db 1741 aggggaacacagataatgtcagacaagaggaaatgctactcgtttcttgatattgagag 1800  
QY 2071 aacgcgaagctgtaacctacacagagaatatacaacgcttctccccaatccagcttggagt 2130  
Db 1801 aacgcgaagctgtaacctacacagagaatatacaacgcttctccccaatccagcttggagt 1860  
QY 2131 cagcttgagagatcagaagttccaagctccaagatcatatgacagacatcaatgctatgct 2190  
Db 1861 cagcttgagagatcagaagttccaagctccaagatcatatgacagacatcaatgctatgct 1920  
QY 2191 ttgagatgcttgagctgtgctgagctgtttgcatgaggtggtcagacatggtacattcaagc 2250  
Db 1921 ttgagatgcttgagctgtgctgagctgtttgcatgaggtggtcagacatggtacattcaagc 1980  
QY 2251 attggaacacagactgactcctctctgctcttctctctgagataactccaacaacaa 2310  
Db 1981 attggaacacagactgactcctctctgctcttctctctgagataactccaacaacaa 2040  
QY 2311 atggtctatagaacacacacccaccctatctccaatctcagagagaacatgcttctcatgctg 2370  
Db 2041 atggtctatagaacacacacccaccctatctccaatctcagagagaacatgcttctcatgctg 2100  
QY 2371 atggaacacacagactgactgagcttctgggtgtgcacaacccagacattctggaacagagc 2430  
Db 2101 atggaacacacagactgactgagcttctgggtgtgcacaacccagacattctggaacagagc 2160  
QY 2431 atgacgccttactgaaggtttctctagttgtgacaagaacacatggttattatltacgagagc 2490  
Db 2161 atgacgccttactgaaggtttctctagttgtgacaagaacacatggttattatltacgagagc 2220  
QY 2491 agttatagaagatatttcagatattctgctgtagtaaaacaatgctcattgaaacaagaagc 2550  
Db 2221 agttatagaagatatttcagatattctgctgtagtaaaacaatgctcattgaaacaagaagc 2280  
QY 2551 ttctcccaagatccaagacacccctagcaactagcaaaagaatcttaatgacccccacca 2610  
Db 2281 ttctcccaagacacacccctagcaactagcaaaagaatcttaatgacccccacca 2298  
QY 2611 gtcttgaacgcacatcaacgggaataaactcgttactactcttcagatcagaatcaagaagaa 2670  
|||||

Db 2299 gtcttgaacgcacatcaacgggaataaactcgttactactcttcagatcagaatcaagaagaa 2358  
QY 2671 attgactatgatataatccatatcaatgattgaaatgaagaagaatatttgacattatgat 2730  
Db 2359 attgactatgatataatccatatcaatgattgaaatgaagaagaatatttgacattatgat 2418  
QY 2731 gaggatgaatatcagaagccccccagcttccaagaagaacacagacacttttatgtc 2790  
Db 2419 gaggatgaatatcagaagccccccagcttccaagaagaacacagacacttttatgtc 2478  
QY 2791 gcaagtgaagagctctggtattatggagatgagtagctccccaatgltctaaagaacag 2850  
Db 2479 gcaagtgaagagctctggtattatggagatgagtagctccccaatgltctaaagaacag 2538  
QY 2851 gctcagagtgagcagtgctccctcagttccaagaagtggtttccaggaatttactgtgc 2910  
Db 2539 gctcagagtgagcagtgctccctcagttccaagaagtggtttccaggaatttactgtgc 2598  
QY 2911 tcttctactcagcccttatccgttggaagaactaatgaaacatttggaactcctgggcca 2970  
Db 2599 tcttctactcagcccttatccgttggaagaactaatgaaacatttggaactcctgggcca 2658  
QY 2971 tatataagacagaagttgaaagataatlatcatggtlaacttccagaaatcagagcctcgt 3030  
Db 2659 tatataagacagaagttgaaagataatlatcatggtlaacttccagaaatcagagcctcgt 2718  
QY 3031 cccctattccttctatctctagcttattcttatgaggaagatcagaagacagaagaa 3090  
Db 2719 cccctattccttctatcttctcccatatcatgaaagaaatcagaagacagaagaa 2778  
QY 3091 cctagaanaaacttctcagagctaatgaaaccaaaacttactttggaagtgacaacat 3150  
Db 2779 cctagaanaaacttctcagagctaatgaaaccaaaacttactttggaagtgacaacat 2838  
QY 3151 cataatgcaaccacataaagaatgagttgagctgcaagcctggtctattctctatggt 3210  
Db 2839 cataatgcaaccacataaagaatgagttgagctgcaagcctggtctattctctatggt 2898  
QY 3211 gacctggaanaaagtgtgcactcagagctgattggaaccccttctgtctgcacactaac 3270  
Db 2899 gacctggaanaaagtgtgcactcagagctgattggaaccccttctgtctgcacactaac 2958  
QY 3271 acaactgaacctgtcatgtggaagacagtgacagtacaagaattctgctgatttccac 3330  
Db 2959 acaactgaacctgtcatgtggaagacagtgacagtacaagaattctgctgatttccac 3018  
QY 3331 atcttctgagacccaagaagctggtactctcaactgaaataatggaagaacatgacaggtc 3390  
Db 3019 atcttctgagacccaagaagctggtactctcaactgaaataatggaagaacatgacaggtc 3078  
QY 3391 cccctgcaatatccagatggaagatcccaacttttaagaagaatatatcgcttccatgaatc 3450  
Db 3079 cccctgcaatatccagatggaagatcccaacttttaagaagaatatatcgcttccatgaatc 3138  
QY 3451 aatggtctacataatgatacactgactggttagtaatggtcagaagatacaagaatctga 3510  
Db 3139 aatggtctacataatgatacactgactggttagtaatggtcagaagatacaagaatctga 3198  
QY 3511 tggatctgctcagcagctgggcagacaaatgaaacatccatctatctaatctcagtgagat 3570  
Db 3199 tggatctgctcagcagctgggcagacaaatgaaacatccatctatctaatctcagtgagat 3258  
QY 3571 ggtgttaactgttaagaaaaaagaggtataaaatggaactgtgaacatctctatccaagt 3630  
Db 3259 ggtgttaactgttaagaaaaaagaggtataaaatggaactgtgaacatctctatccaagt 3318  
QY 3631 gttttgagaacagtgaanaatgttaccatccaagcctggaatttgggggtgagatgcctt 3690  
Db 3319 gttttgagaacagtgaanaatgttaccatccaagcctggaatttgggggtgagatgcctt 3378  
QY 3691 attgagcagatctacatgctggtgagacacactttctggtgtacagcaataaagtgt 3750  
Db 3379 attgagcagatctacatgctggtgagacacactttctggtgtacagcaataaagtgt 3438



QY	3751	cgactcccttgggaatgagctctcttgagacaatatgagatcttcagatatacagcttcagga	3810
Db	3439	cgactcccttgggaatgagctctcttgagacaatatgagatcttcagatatacagcttcagga	3498
QY	3811	caatatgagacagtgaggcccaaaagcttgagcagacttcatattccggatcaatacagcc	3870
Db	3499	caatatgagacagtgaggcccaaaagcttgagcagacttcatattccggatcaatacagcc	3558
QY	3871	tgaggacccaagagagccctttctcttgatcaaaagtggatcgtcttgagaccaatgatatt	3930
Db	3559	tgaggacccaagagagccctttctcttgatcaaaagtggatcgtcttgagaccaatgatatt	3618
QY	3931	caagcgatcaagaagccagggctgcggtcagaagtcttcacagctctacatctcagttc	3990
Db	3619	caagcgatcaagaagccagggctgcggtcagaagtcttcacagctctctacatctcgaatt	3678
QY	3991	atcatcatgtatgctcttgaatggaaaagaatggagaacttatcgagaatcttcacagga	4050
Db	3679	atcatcatgtatgctcttgaatggaaaagaatggagaacttatcgagaatcttcacagga	3738
QY	4051	accttaatgctctctcttgacaatgttgatgatactcttgagataaaacaatatctttac	4110
Db	3739	acctcatgctctctcttttgacaatgttgatgatactcttgagataaaacaatatctttac	3798
QY	4111	cctcccaattatgtctcgatatacatccggtttgacacccaactaatatagatctgcagact	4170
Db	3799	cctcccaattatgtctcgatatacatccggtttgacacccaactaatatagatctgcagact	3858
QY	4171	cttcgcatgtagagtgtgtagctcttgatataaagtgtgcagacatcttgagaaatggag	4230
Db	3859	cttcgcatgtagagtgtgtagctcttgatataaagtgtgcagacatcttgagaaatggag	3918
QY	4231	agtaaaagcaatatcagatgacacagatctactgcttcaactcttaccaatatgtttgcc	4290
Db	3919	agtaaaagcaatatcagatgacacagatctactgcttcaactcttaccaatatgtttgcc	3978
QY	4291	acctggtctcccttccaaaagctctgacttcaactcccaaggagagtaatgctctgagact	4350
Db	3979	acctggtctcccttccaaaagctctgacttcaactcccaaggagagtaatgctctgagact	4038
QY	4351	cagggttaataatccaaaagaatggtctgcgaagtggactctcagaagaacatgaaagtaca	4410
Db	4039	cagggttaataatccaaaagaatggtctgcgaagtggactctcagaagaacatgaaagtaca	4098
QY	4411	ggagtaactactcaggagagtaaaatcctctgtcttaccacagatgtatgtgaaagagttctc	4470
Db	4099	ggagtaactactcaggagagtaaaatcctctgtcttaccacagatgtatgtgaaagagttctc	4158
QY	4471	atctccagcagttcaagaatgagtcacatcagttgactctctttttccaagaatgycaaaagttaag	4530
Db	4159	atctccagcagttcaagaatgagtcacatcagttgactctctttttccaagaatgagaaataaa	4218
QY	4531	gtttttcagggaataataagaatctcttcaaacctgtgtgtgaacctcttagcccaacgtta	4590
Db	4219	gtttttcagggaataataagaatctcttcaaacctgtgtgtgaacctcttagcccaacgtta	4278
QY	4591	ctgactcgtactaccttcgaattcaacccccagagtttgggtgacacagatctgcctgagagt	4650
Db	4279	ctcactcgtactaccttcgaattcaacccccagagtttgggtgacacagatctgcctgagagt	4338
QY	4651	gaggttcttggtctgtcgagagcaagaactctactg	4685
Db	4339	gaggttcttggtctgtcgagagcaagaactctactg	4373

RESULT	15
AA081544	
ID	AA081544 standard; DNA; 4830 BP
XX	
AC	AA081544;
XX	
DT	04-DEC-1990 (first entry)

XX Human Factor VIII-C analog having exon 14 deleted.  
DE  
XX  
XX Human Factor VIII-C analog; exon deletion; coagulation disorders;  
KW haemophilia; ss.  
XX  
XX EP265778-A.  
XX  
XX 04-MAY-1988.  
PD  
XX 14-OCT-1987; 87EP-0115043.  
XX  
XX 15-OCT-1986; 86US-0919153.  
PR  
XX  
XX (RORE ) RORER INT OVERSEAS.  
PA  
XX  
XX Sarver N, Drohan W;  
PI  
XX  
XX WPI; 1988-120930/18.  
DR  
XX  
XX Human Factor VIII-C analogue free of other proteins - produced from  
PT recombinant DNA and used in the treatment of coagulation disorders.  
XX  
XX  
XX Claim 3; Page 13; 42pp; English.  
PS  
XX  
XX The genetically engineered analogue can provide a dependable and  
CC readily available therapeutic agent to be used in the treatment of  
CC haemophilia and coagulation disorders in humans or animals.  
CC See also AAN81543-45.  
CC  
XX  
XX Sequence 4830 BP; 1392 A; 1091 C; 1058 G; 1289 T; 0 other;  
XX

Query Match	87.1%;	Score 4152.2;	DB 9;	Length 4830;
Best Local Similarity	94.8%;	Pred. No. 0;		
Matches 4386;	Conservative	0;	Mismatches 43;	Indels 198; Gaps 1
QY 271	atgaaatagagcttcacacgtgctctcttccttgctgcttgagatctgcttagtgc	330		
Db 1	atgcaaatagagcttcacacgtgctctcttccttgctgcttgagatctgcttagtgc	60		
QY 331	accgaagaatactactcctbgtggtgagtgagactgtcatatggaatatalgcaagtac	390		
Db 61	accgaagaatactactcctbgtggtgagtgagactgtcatatggaatatalgcaagtac	120		
QY 391	ggtgagctgcctctgtgagcgcaagaattcctcctbtagatgtgccaaaatctttccat	450		
Db 121	ggtgagctgcctctgtgagcgcaagaattcctcctbtagatgtgccaaaatctttccat	180		
QY 451	accacagctgtgtacaaaagaactcgtttgtagatcacaggtcacactttcaacac	510		
Db 181	accacagctgtgtacaaaagaactcgtttgtagatcacaggtcacactttcaacac	240		
QY 511	gctaaagcgaagcgaccctctgtatgtgtctgtctagagtcctacacacacaggtcga	570		
Db 241	gctaaagcgaagcgaccctctgtatgtgtctgtctagagtcctacacacacaggtcga	300		
QY 571	gatacagtggtcatctaaccttaagaacatgcttccactcctgtcagttcatgtggt	630		
Db 301	gatacagtggtcatctaaccttaagaacatgcttccactcctgtcagttcatgtggt	360		
QY 631	ggtgtatcccaactcgtgaaagctctcttagagagagctgaatatgtatgtacgacag	690		
Db 361	ggtgtatcccaactcgtgaaagctctcttagagagagctgaatatgtatgtacgacag	420		
QY 691	gagaaagaatgatataaagctctccctgtgtgaaagccataatattgtctgacagtc	750		
Db 421	gagaaagaatgatataaagctctccctgtgtgaaagccataatattgtctgacagtc	480		
QY 751	aaagagaatgttccaatgtgcctctgaccacatgtgcttactcatcatatctttcat	810		
Db 481	aaagagaatgttccaatgtgcctctgaccacatgtgcttactcatcatatctttcat	540		



QY 811 gtggacctggtlaaaaagcttgaaltcagggcccatctgagccctactagatgtagaagaa 870  
 Db 541 gtggacctgtlaaaaagcttgaaltcagggcccatctgagccctactagatgtagaagaa 600  
 QY 871 gggagctctggccaagggaaaagacacagaccttgcacaatttatactacttttgcgtga 930  
 Db 601 gggagctctggccaagggaaaagacacagaccttgcacaatttatactacttttgcgtga 660  
 QY 931 ttgatataaggaaaagtctggacctcagaaacaaagaccttgcagggtaaggat 990  
 Db 661 ttgatataaggaaaagtctggacctcagaaacaaagaccttgcagggtaaggat 720  
 QY 991 gctgcactctcctgggaccttggcctlaaaatgcacacagtcactgtatgtfaaacaggtct 1050  
 Db 721 gctgcactctcctgggaccttggcctlaaaatgcacacagtcactgtatgtfaaacaggtct 780  
 QY 1051 ctgcagagctctgatatgtgacacaggaanaatcagttactatggcactgtatgtatggc 1110  
 Db 781 ctgcagagctctgatatgtgacacaggaanaatcagttactatggcactgtatgtatggc 840  
 QY 1111 aacctcctgaagtgcacataatctcctgaaggtgcacacatttctgtgaggaacacat 1170  
 Db 841 aacctcctgaaggtgcacataatctcctgaaggtgcacacatttctgtgaggaacacat 900  
 QY 1171 cgcgaagcgctccttggaatctgcgcacataacttcccttactcgtcacaacaccttgatg 1230  
 Db 901 cgcgaagcgctccttggaatctgcgcacataacttcccttactcgtcacaacaccttgatg 960  
 QY 1231 gaaccttgagacagtttactactgttttgcatactcttccacacacatgacttgagaa 1290  
 Db 961 gaaccttgagacagtttactactgttttgcatactcttccacacacatgacttgagaa 1020  
 QY 1291 gcttatgtcaaaatgaagaacgtctccagaggaaagcccaactagagaatgaanaataatgaa 1350  
 Db 1021 gcttatgtcaaaatgaagaacgtctccagaggaaagcccaactagagaatgaanaataatgaa 1080  
 QY 1351 gaagcggaagaactatgatgtatgtactactgacttgcgaatgtgactgtgcagagttgat 1410  
 Db 1081 gaagcggaagaactatgatgtatgtactactgacttgcgaatgtgactgtgcagagttgat 1140  
 QY 1411 gatgacaacctctcctcttataccaaatcgtcagttcagttgcagaagaaccttaaaact 1470  
 Db 1141 gatgacaacctctcctcttataccaaatcgtcagttcagttgcagaagaaccttaaaact 1200  
 QY 1471 ttgggtataatfatcatgtcgtctggaaggaagagacttggagacttcccttaagccctcgcc 1530  
 Db 1201 ttgggtataatfatcatgtcgtctggaaggaagagacttggagacttcccttaagccctcgcc 1260  
 QY 1531 cccgatatgacagaagtataaaaagtcacatatttgaacatgtgcccctcagcgagttgtag 1590  
 Db 1261 cccgatatgacagaagtataaaaagtcacatatttgaacatgtgcccctcagcgagttgtag 1320  
 QY 1591 aagttacaaaaaagtcogattatgtgcatacacagatgaaacctttaaagctgtgaagct 1650  
 Db 1321 aagttacaaaaaagtcogattatgtgcatacacagatgaaacctttaaagctgtgaagct 1380  
 QY 1651 attcagaataaattcagaagaactctggagaccttacttattatggaggaagtgtgagacacatg 1710  
 Db 1381 attcagaataaattcagaagaactctggagaccttacttattatggaggaagtgtgagacacatg 1440  
 QY 1711 ttgatataatgaagaatcagaagacagacacatalaacacttaccctcagcgatcaact 1770  
 Db 1441 ttgatataatgaagaatcagaagacagacacatalaacacttaccctcagcgatcaact 1500  
 QY 1771 gatgtccgtctcttgcattcaagagagattaccaaaggtgtlaaacaatttgaagatttt 1830  
 Db 1501 gatgtccgtctcttgcattcaagagagattaccaaaggtgtlaaacaatttgaagatttt 1560  
 QY 1831 ccaattctgcagaggaanaaatcacaatataaatgagacagtgactgtagaagaattggcca 1890  
 Db 1561 ccaattctgcagaggaanaaatcacaatataaatgagacagtgactgtagaagaattggcca 1620  
 QY 1891 actaaatcagatcctcgtgtcgctgaccggtatctactctgattcglttaatatgagaga 1950

Db 1621 actaaatcagatcctcgtgtcgctgaccggtatctactctgattcglttaatatgagaga 1680  
 QY 1951 gatctagcttcaagactcatcttgccctctccatctgtctcaaaaagaatctgatatcaa 2010  
 Db 1681 gatctagcttcaagactcatcttgccctctccatctgtctcaaaaagaatctgatatcaa 1740  
 QY 2011 agaggaacacagataagttacagaagaagaatgtcaatccctgtttctgtatgtatgag 2070  
 Db 1741 agaggaacacagataagttacagaagaagaatgtcaatccctgtttctgtatgtatgag 1800  
 QY 2071 aacccgaagctgtgactcctcacaaggaatatalacaacgctttctcccaatccagcttgagtg 2130  
 Db 1801 aacccgaagctgtgactcctcacaaggaatatalacaacgctttctcccaatccagcttgagtg 1860  
 QY 2131 cagcttgaggaatccagaagttccaaagctccacaatcatgtcagaagcatcaatgtgtatgt 2190  
 Db 1861 cagcttgaggaatccagaagttccaaagctccacaatcatgtcagaagcatcaatgtgtatgt 1920  
 QY 2191 ttgatagtttggaagttggaagttgttgcataaggttggcagactgtgacatcttaagc 2250  
 Db 1921 ttgatagtttggaagttggaagttgttgcataaggttggcagactgtgacatcttaagc 1980  
 QY 2251 attggaacacagactgactcctctctgtcttctctctgtgatatataccttcaacaacaa 2310  
 Db 1981 attggaacacagactgactcctctctgtcttctctctgtgatatataccttcaacaacaa 2040  
 QY 2311 atgtgtctatgaagaacacacactacactatcccaattcctcagaagaacactgtcttcatgtcg 2370  
 Db 2041 atgtgtctatgaagaacacacactacactatcccaattcctcagaagaacactgtcttcatgtcg 2100  
 QY 2371 atggaanaacccaaggtctatgtatgtctgggtgtgcacaacactcagaacttctggaacaagagc 2430  
 Db 2101 atggaanaacccaaggtctatgtatgtctgggtgtgcacaacactcagaacttctggaacaagagc 2160  
 QY 2431 atgacggccttacttgaaggttcttctagtttgcacaagaacacttgcattatctacggagac 2490  
 Db 2161 atgacggccttacttgaaggttcttctagtttgcacaagaacacttgcattatctacggagac 2220  
 QY 2491 agttatgaagaatattccagcatactctgtcgtgatraaaaaatgacatgtgacaagaagc 2550  
 Db 2221 agttatgaagaatattccagcatactctgtcgtgatraaaaaatgacatgtgacaagaagc 2280  
 QY 2551 ttctccagaatctca----- 2655  
 Db 2281 ttctccagaatctcaatctatgtgactctctgtctgtgataaacatagtgtactcagata 2340  
 QY 2655 ----- 2565  
 Db 2341 ccaaaaagaagagtggaatccccaagaaggtcaaccagaanaaacagctttaaagaanaag 2400  
 QY 2565 ----- 2565  
 Db 2401 gatcacatttgcctcctgaacgcttctgaaagcaatcatgtcaatagcaacaatagag 2460  
 QY 2566 ----- agcacacctgacactgtgccaagaagcaa 2592  
 Db 2461 ggaacaaataagccgaaatagaagtcacacctggtgcacgaagaagtgagactgaaagctgt 2520  
 QY 2593 tttaatgcccaccccaacagttcttgaagaacgcatcaacgaggaataactcgttactactct 2652  
 Db 2521 tgtctccaaaacccacacagttcttgaagaacgcatcaacgaggaataactcgttactactct 2580  
 QY 2653 cagtcagatcaagaaggaatctgactatgtatgataccatatcagttgaaatgaagaagaa 2712  
 Db 2581 cagtcagatcaagaaggaatctgactatgtatgataccatatcagttgaaatgaagaagaa 2640  
 QY 2713 gatttgacaatttatatagaagttgaanaatcagagcccgagcgtttaaagaagaanaa 2772  
 Db 2641 gatttgcacattatagatgagatgaaatcagagcccgagcgtttaaagaagaanaa 2700  
 QY 2773 cgaactatttatactgtcagtgagagagccttgagattatagatgagtagctccca 2832



Db 2701 cgacactatttlatgtctgagctggaagctctggaattatggatagtagctccca 2760  
QY 2833 catgtcttaagaaacagggtccagatggtgcctcagttcaagaagtgttttc 2892  
Db 2761 catgtcttaagaaacagggtccagatggtgcctcagttcaagaagtgttttc 2820  
QY 2893 caggaatttactgtatgctccttactcagccctataccgtggaagaaataaagaat 2952  
Db 2821 caggaatttactgtatgctccttactcagccctataccgtggaagaaataaagaat 2880  
QY 2953 ttggagctccggggccaataataagacagaagtgaagaataataatgttaatttc 3012  
Db 2881 ttggagctccggggccaataataagacagaagtgaagaataataatgttaatttc 2940  
QY 3013 agaatacaggcctctcgtccctatcctctatctatagcctatttcttatgagaagat 3072  
Db 2941 agaatacaggcctctcgtccctatcctctatctatagcctatttcttatgagaagat 3000  
QY 3073 cagagcagaaggacagaacctagaaaaaacttgcagcctaataagaaacaaacttac 3132  
Db 3001 cagagcagaaggacagaacctagaaaaaacttgcagcctaataagaaacaaacttac 3060  
QY 3133 ttcttggaagtgcacaactcatatagtcaccctaagaatgagtgtgactgcaaaagctcg 3192  
Db 3061 ttcttggaagtgcacaactcatatagtcaccctaagaatgagtgtgactgcaaaagctcg 3120  
QY 3193 gcttatttctctgaatgttgacctgaaaaaagtgtgcactcaggtctgattggaacctt 3252  
Db 3121 gcttatttctctgaatgttgacctgaaaaaagtgtgcactcaggtctgattggaacctt 3180  
QY 3253 ctgtgtctgcacaactaaacacatgaaacctgtcatgaggaacaagtgcagttacaggaa 3312  
Db 3181 ctgtgtctgcacaactaaacacatgaaacctgtcatgaggaacaagtgcagttacaggaa 3240  
QY 3313 ttgtgctgtgttttccacatcttggatgagaccaaagctgtgacttaccgaaaaattg 3372  
Db 3241 ttgtgctgtgttttccacatcttggatgagaccaaagctgtgacttaccgaaaaattg 3300  
QY 3373 gaaagaacatgcagaggtccctcgtcaataataccagaatggaagaaacccactttaagaagat 3432  
Db 3301 gaaagaacatgcagaggtccctcgtcaataataccagaatggaagaaacccactttaagaagat 3360  
QY 3433 tatcgtctcattcgaatccaatggtctacataatggaatacacactcgtgttaattgct 3492  
Db 3361 tatcgtctcattcgaatccaatggtctacataatggaatacacactcgtgttaattgct 3420  
QY 3493 caggaatcaaggatcgtatgctatcgtcagcatggtgcagcaaatgaaacatccattct 3552  
Db 3421 caggaatcaaggatcgtatgctatcgtcagcatggtgcagcaaatgaaacatccattct 3480  
QY 3553 attcatttcagtgagacatgtgtcactgtacgaaaaaagaagagataaaaaatgacatg 3612  
Db 3481 attcatttcagtgagacatgtgtcactgtacgaaaaaagaagagataaaaaatgacatg 3540  
QY 3613 tacaattctcatccaggtgtttttggaacagtggaatggttacatccaagctggaaat 3672  
Db 3541 tacaattctcatccaggtgtttttggaacagtggaatggttacatccaagctggaaat 3600  
QY 3673 tggcgggtggaatgccttatgtggcgaacatctacatcctggatcgagacactttttctg 3732  
Db 3601 tggcgggtggaatgccttatgtggcgaacatctacatcctggatcgagacactttttctg 3660  
QY 3733 gtgtacagcaataagtgtagactccctgtggaatggtctcttgacacatatagatttt 3792  
Db 3661 gtgtacagcaataagtgtagactccctgtggaatggtctcttgacacatatagatttt 3720  
QY 3793 cagattaacagcttcaggaacaatatgacaagtggcccaagctggtccagacttcatat 3852  
Db 3721 cagattaacagcttcaggaacaatatgacaagtggcccaagctggtccagacttcatat 3780  
QY 3853 tccggatcaatcaatgctctgagacacaagaagccttttcttgatcaagaatgtagatg 3912  
Db 3781 tccggatcaatcaatgctctgagacacaagaagccttttcttgatcaagaatgtagatg 3840

QY 3913 ttggccaatgatattatcacgcatcaagaaccagggtgccgttcagaaattctccagc 3972  
Db 3841 ttggccaatgatattatcacgcatcaagaaccagggtgccgttcagaaattctccagc 3900  
QY 3973 ctctacatctctcagttatcatcatatgataagcttgatggaagaaatggtcagacttat 4032  
Db 3901 ctctacatctctcagttatcatcatatgataagcttgatggaagaaatggtcagacttat 3960  
QY 4033 cgaagaaattccactcggaaacttaatggtcttcttttgcaatggtgattatctgggata 4092  
Db 3961 cgaagaaattccactcggaaacttaatggtcttcttttgcaatggtgattatctgggata 4020  
QY 4093 aaacacaatttttaaccctccaatttatgtctgataacatccgtttgcaaccaactcat 4152  
Db 4021 aaacacaatttttaaccctccaatttatgtctgataacatccgtttgcaaccaactcat 4080  
QY 4153 tatagacttcgacgacctcttcgcaatggaatgtagatggtgtgatttaaatagttgacg 4212  
Db 4081 tatagacttcgacgacctcttcgcaatggaatgtagatggtgtgatttaaatagttgacg 4140  
QY 4213 atgcattgggaatgggaagttaagcaatatcagatgcacagattactgtcttaacctac 4272  
Db 4141 atgcattgggaatgggaagttaagcaatatcagatgcacagattactgtcttaacctac 4200  
QY 4273 ttaccaaataatgttgcacacctgttctccttcaaaaagctcgacttcaaccccaaggagg 4332  
Db 4201 ttaccaaataatgttgcacacctgttctccttcaaaaagctcgacttcaaccccaaggagg 4260  
QY 4333 agtaatgcttgagacactcaggtgaataataccaaaagatggtctgcaagtgcacttcag 4392  
Db 4261 agtaatgcttgagacactcaggtgaataataccaaaagatggtctgcaagtgcacttcag 4320  
QY 4393 aagaacaatggaatgcacaggagttaactctcagggagttaaatctcgtcttaccagatg 4452  
Db 4321 aagaacaatggaatgcacaggagttaactctcagggagttaaatctcgtcttaccagatg 4380  
QY 4453 tatgtgaagaggttctctcatctccacagcatgcaaatggtgcactcgtcttctttt 4512  
Db 4381 tatgtgaagaggttctctcatctccacagcatgcaaatggtgcactcgtcttctttt 4440  
QY 4513 cagaatggtcagaagttaaggttttccagggaatacaagactcctcacacctgtgtgaaac 4572  
Db 4441 cagaatggtcagaagttaaggttttccagggaatacaagactcctcacacctgtgtgaaac 4500  
QY 4573 tctctagaagccacccgttaactgactcgtctacacttgaattcaaccccaaggttgggtgac 4632  
Db 4501 tctctagaagccacccgttaactgactcgtctacacttgaattcaaccccaaggttgggtgac 4560  
QY 4633 cagattgcccgtgaagatggaagtctctggtcgtgcaagcacagagcctctactagagggtg 4692  
Db 4561 cagattgcccgtgaagatggaagtctctggtcgtgcaagcacagagcctctactagagggtg 4620  
QY 4693 cgaagtgc 4699  
Db 4621 ccaactgac 4627

Search completed: November 18, 2001, 05:25:05  
Job time: 37037 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 18, 2001, 01:27:15 ; Search time 22045.3 Seconds  
(Without alignments)  
3343.294 Million cell updates/sec

Title: US-09-689-430-1\_COPY\_150\_4914  
Perfect score: 4765  
Sequence: 1 cctcttcaagtaaacagta.....ttgggtcgttctgttcgagtc 4765

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Maximum Match 0%  
Listing first 45 summaries

Database :

GenEmbl: \*  
1: gb\_ba1: \*  
2: gb\_ba2: \*  
3: gb\_ba3: \*  
4: gb\_in1: \*  
5: gb\_in2: \*  
6: gb\_in3: \*  
7: gb\_om: \*  
8: gb\_ov: \*  
9: gb\_pat1: \*  
10: gb\_pat2: \*  
11: gb\_ph: \*  
12: gb\_pl1: \*  
13: gb\_pl2: \*  
14: gb\_pl3: \*  
15: gb\_pl4: \*  
16: em\_ba1: \*  
17: em\_ba2: \*  
18: em\_fun: \*  
19: em\_htgo\_hum: \*  
20: em\_htgo\_inv: \*  
21: em\_htgo\_rod: \*  
22: em\_htg\_hum1: \*  
23: em\_htg\_hum2: \*  
24: em\_htg\_hum3: \*  
25: em\_htg\_hum4: \*  
26: em\_htg\_hum5: \*  
27: em\_htg\_hum6: \*  
28: em\_htg\_hum7: \*  
29: em\_htg\_hum8: \*  
30: em\_htg\_inv1: \*  
31: em\_htg\_inv2: \*  
32: em\_htg\_other: \*  
33: em\_htg\_rod: \*  
34: em\_hum1: \*  
35: em\_hum2: \*  
36: em\_hum3: \*  
37: em\_hum4: \*  
38: em\_hum5: \*  
39: em\_hum6: \*  
40: em\_hum7: \*  
41: em\_in: \*  
42: em\_om: \*  
43: em\_or: \*

150-4835

44: em\_ov: \*  
45: em\_pat: \*  
46: em\_ph: \*  
47: em\_pl: \*  
48: em\_ro: \*  
49: em\_sts: \*  
50: em\_sy: \*  
51: em\_un: \*  
52: em\_vi: \*  
53: gb\_sts1: \*  
54: gb\_sts2: \*  
55: gb\_sts3: \*  
56: gb\_sy: \*  
57: gb\_un: \*  
58: gb\_v11: \*  
59: gb\_v12: \*  
60: gb\_htg1: \*  
61: gb\_htg2: \*  
62: gb\_htg3: \*  
63: gb\_htg4: \*  
64: gb\_htg5: \*  
65: gb\_htg6: \*  
66: gb\_htg7: \*  
67: gb\_htg8: \*  
68: gb\_htg9: \*  
69: gb\_htg10: \*  
70: gb\_htg11: \*  
71: gb\_htg12: \*  
72: gb\_htg13: \*  
73: gb\_htg14: \*  
74: gb\_htg15: \*  
75: gb\_htg16: \*  
76: gb\_htg17: \*  
77: gb\_htg18: \*  
78: gb\_htg19: \*  
79: gb\_htg20: \*  
80: gb\_htg21: \*  
81: gb\_htg22: \*  
82: gb\_htg23: \*  
83: gb\_htg24: \*  
84: gb\_htg25: \*  
85: gb\_pr1: \*  
86: gb\_pr2: \*  
87: gb\_pr3: \*  
88: gb\_pr4: \*  
89: gb\_pr5: \*  
90: gb\_pr6: \*  
91: gb\_pr7: \*  
92: gb\_pr8: \*  
93: gb\_pr9: \*  
94: gb\_rot1: \*  
95: gb\_rot2: \*  
96: gb\_in4: \*  
97: gb\_pr10: \*  
98: em\_ba3: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	4411.6	92.6	9354	9 AR003585	AR003585 Sequence
2	4357.6	91.5	4670	9 AR110040	AR110040 Sequence
3	4148.8	87.1	4278	10 I08644	I08644 Sequence 4
4	4147.2	87.0	4281	10 I08643	I08643 Sequence 3
5	4085	85.7	4548	10 I08642	I08642 Sequence 2
6	4085	85.7	4551	10 I08641	I08641 Sequence 1
7	3748	78.7	5035	9 AR034084	AR034084 Sequence
8	3748	78.7	5035	9 AR071306	AR071306 Sequence



9	3189.8	66.9	433.4	9	AR029098
10	2605.4	54.7	445.1	9	AR110041
111	2336.4	49.0	9009	9	AR003710
112	2336.4	49.0	9009	9	AR029065
113	2336.4	49.0	9009	10	IJ1901
114	2336.4	49.0	9009	10	IG3424
115	2336.4	49.0	9009	97	HNEVITII
116	2335.4	49.0	7056	10	IJ27063
117	2334.8	49.0	8241	9	A05328
118	2334.8	49.0	8241	9	A07042
119	2334.8	49.0	8831	9	E00527
120	2334.8	49.0	9029	97	HNEVITIIIC
121	2333.8	49.0	7272	10	I05404
122	2333.2	49.0	8667	9	AX052730
123	2333.2	49.0	8967	10	I17409
124	2333.2	49.0	8967	93	HSEVIIR
125	2328.4	48.9	7440	10	I08345
126	2314	48.6	7440	10	I08457
127	2275.2	47.7	6999	10	I17105
128	1848	38.8	7032	7	AF016234
129	1840	38.6	7415	7	AF049489
130	1802.8	37.8	7433	9	AR003712
131	1802.8	37.8	7433	9	AR029067
132	1802.8	37.8	7433	94	I63427
133	1802.8	37.8	7433	94	M0SCCVIIII
134	1752.2	36.8	6539	7	SS049517
135	1751.2	36.8	6402	9	AR029097
136	1661	34.9	1993	10	I02047
137	1603.2	33.6	1728	10	I02054
138	1240.4	26.0	1623	10	I27064
139	991.4	20.8	3852	10	I04400
140	990.8	20.8	3852	10	E00422
141	893.4	18.7	1130	9	AR003711
142	893.4	18.7	1130	9	AR029066
143	893.4	18.7	1130	10	I13190
144	893.4	18.7	1130	10	I63426
145	675.4	14.2	955	10	E00526
					AR029098 Sequence
					AR110041 Sequence
					AR003710 Sequence
					AR029065 Sequence
					IJ1901 Sequence 3
					IG3424 Sequence 1
					K01240 Human coagu
					I27063 Sequence 1
					A05328 Synthetic F
					A07042 Artificial
					E00527 Complete cd
					M4113 Human coagu
					I05404 Sequence 3
					AX052730 Sequence
					I17409 Sequence 1
					X01179 Human mRNA
					I08345 Sequence 2
					I08457 Sequence 1
					I17105 Sequence 1
					AF016234 Canis fam
					AF049489 Canis fam
					AR003712 Sequence
					AR029067 Sequence
					I63427 Sequence 7
					L05573 Mus domestil
					O49517 Sus scrofa
					AR029097 Sequence
					I02047 Sequence 1
					I02054 Sequence 8
					I27064 Sequence 3
					I04400 Sequence 35
					E00422 Human genom
					AR003711 Sequence
					AR029066 Sequence
					I13190 Sequence 5
					I63426 Sequence 1
					E00526 Part of pES

## ALIGNMENTS

OY	326	gtgccaccacaaagaatactactctggatgcagatggaactgcatacggacataatgcaaaatg	385
Db	3020	GTGCCACCAAGATACTACTCTGGGTGCAATGGACACTGCATGGACATATATGCAAAAGTG	3079
OY	386	atccctggatgagctctccctgctgagacgcaaaatcttcctctctagatgtgccaaatctttccat	445
Db	3080	ATCTCGGTGAGCTCCGTGTGGACGAAGATTCTCTGTAGTGGCCAAATCTTTTCAT	3139
OY	446	tcaaacctccagtcgctgtatcaaaaagactctgtttgtagaatcaacgltcaacttca	505
Db	3140	TCAACACCTCAAGTGTGTACAAAAAGACTCTGTTTGTAGAAATTCACGGTTCCACCTTTTCA	3189
OY	506	acatcgcctaagccaagggccacccttgatggtctctgtctagtcctcaatccaaggtctgg	565
Db	3200	ACATCGGTAAAGCCAAAGGCCACCTCGTAGTGGTGTGTAGGTCTCTACACCAACGGCTGAGG	3259
OY	566	tttatgtatcagctgtgcatatacctaagaagaactggtcccatccctgtcgaattctcatg	625
Db	3260	TTTATGTATCAGTGTGTATTTACCTTTAAGACATGGCTTCCATCTGTCACTCTTTCATG	3319
OY	626	ctctgttgatctaccctactgtgaaagctctctgaaggagctgaatatatgatctacgacaagtc	685
Db	3320	CTGTGTGTATCTCTACTGTGGAAGGCTTGGAGGAGCTGAATATGATGTAGACCACTC	3379
OY	686	aaagggagaaagaatgataaaagctctccctcctgtgtgaaagccatacatatgtctgcgaag	745
Db	3380	AAAGGGAGAAAGAGATTAAGATCTTCCCTGTGGAAAGCCATACATATGTCTGGCAGG	3439
OY	746	tcccgaaagaagaatggtccaatgtgctctgcaccatctggtccttaactcaatacttt	805
Db	3440	TCTCGAAAGAGAAATGGTCCATATGGCCCTGTACCCACTGTGCCCTTACTATATCTATCTTT	3499
OY	806	ctcaatgtgagcctgttaaaaagacttgaattcaaguccctcaatggaagccctactagatgtatg	865
Db	3500	CTCATGTGGACACTGTTAAAGACTTGTGAATTCAGAGCCTCATTTGGAGGCCCTACGTAGTATGTA	3559
OY	866	gagaagaggatctcgtgcgaagaagaacagaagccttgacaaaattatbaactatttctg	925

RESULT	1				
AR003585					
LOCUS	AR003585	9354 bp	DNA	PAT	04-DEC-1998
DEFINITION	Sequence 2 from patent US 5744326.				
ACCESSION	AR003585				
VERSION	AR003585.1	GI:3964844			
KEYWORDS	.				
SOURCE	Unknown.				
ORGANISM	Unknown.				
REFERENCE	Unclassified.				
AUTHORS	1 (bases 1 to 9354)				
TITLE	Ill,C.R. and Bidlingmayer,S. Use of viral cis-acting post-transcriptional regulatory sequences to increase expression of intronless genes containing near-consensus splice sites Patent: US 5744326-A 2 28-APR-1998;				
JOURNAL	Location/Qualifiers				
FEATURES	1..9354				
SOURCE	/organism="unknown"				
BASE COUNT	2506 a 2239 c 2161 g 2448 t				
ORIGIN					

Query Match	92.6%;	Score 4411.6;	DB 9;	Length 9354;
Best Local Similarity	99.7%;	Pred. Ns: 0;		
Matches 4420; Conservative	0;	Mismatches 14;	Indels 0;	Gaps 0;
Oy	266	ccacatgcgaatagagctccacactgctctcttgcgcctttgcgattcgttta	325	
Db	2960	ccaccatggtttatagactctccacctgctttcttctgtgccttttgcgattcgttta	3019	

QY	986	gggttgatgatctgcctgcggccttgcccttaaaatgaaacaaatgaatgtaatgtaaa	1045
Db	3680	GGGATGCTGCATCTGCTCGGGGCGCTGGCCCTTAATATGCACACGTCAATGGATTGTAACA	3739
QY	1046	ggtctgcacagatctgattatggaatggcaaggaatcagtcattatggcaatgfatggaa	1105
Db	3740	GGTCTTGCCAGGTTCTGATTGGAATGCCACAGGAATCACTCTATTGGCATGTGATTGGAA	3799
QY	1106	tgaggacaccttcctggaatgagatcaataatcttcgaaagtcaacaattctctgga	1166
Db	3800	TGGGACACACTCTCTGAAGTCACTCAATATCTCTCGAAGTACACATTCTTTGGAGAA	3855
QY	1166	accatcgccagcgctctcttggaatctgcgcaataacttctctactgctcaaacactct	1225
Db	3860	ACCATGCCACAGGGCTCTTGGAATCTCGCATTAATCTTCTTACTGTCAAAACACTCT	3919
QY	1226	tgatgacacttggaacagtttctactglttgtgcatactctctccaacaatgtaggca	1285
Db	3920	TGATGGAACCTTGGACAGTTTCTCACTGTTTGTCAATCTTCCACCAACATATGATGGCA	3979
QY	1286	tggaagcttatgtcoaatgtagacagctgtgcagaggaaaccccaactacgaatgtaaaata	1345
Db	3980	TGGAAGCTTATGTCAAGTATGACAGCTGTGCCAGAGAACCCCAACTACGAATGAATAATA	4039
QY	1346	atgaagaagcgggaagactatgatgatgatacttactgtaattctgaaatgtagtgtgcagt	1405
Db	4040	ATGAAGAAGCGGAAGACTATGATGATGATCTTACTGATTTCTGAATGATGTGGTCACAGT	4099
QY	1406	ttgatgatgaaactctctctcttacttaccaaattgcgtcagttgcccagaagacatccta	1465



```
Db 4100 TTGATGATGACACGCTTCCTTCCTTTATCCAAATTCGCTCAGTTGCCAAGAAGATCCTA 4159
Qy 1466 aaacttggfatacatcaatctgctgctgaagaggagactggactatgcttccttagtcc 1525
Db 4160 AACTTGGGTACATTACATTCCTCTGTAAGAGGAGACTGGACTATGCTCCCTTAGTCC 4219
Qy 1526 tcgccccgatgacagaagaattataaaagtcaatatgtgaacaatggcccttaaggattg 1585
Db 4220 TCGCCCCGATGACAGAAAGTTATAAAGTCAATTTTGAACAAAGGCCCTCAGCGGATGG 4279
Qy 1586 gtaagaagtacaaaagctccgactttagcatatgcatcacagatgaacactttaagactg 1645
Db 4280 GTAGAGAGTACAAAAGTCCGATTTATGCGATACAGATGAAGAACTTTAAGCTGCTG 4339
Qy 1646 aagctatcagcatgaaatcagaatcttggagacttaactttaggggaagttagagaca 1705
Db 4340 AAGCTATTCAGCATGAATTCAGGAATCTTGGGACCTTACCTTATGGGGAAGTTGAGACA 4399
Qy 1706 cactgttatattatattagaatcaagaacagacatatacatctacccctcagaa 1765
Db 4400 CACTGTGATTAATTAATTAAGATCAAGCAAGCAACATTAACATCTACCTCAGCGAA 4459
Qy 1766 tcactgatlccgctccttltatcaaggagattaccaaaagtgtaaaacatttgaag 1825
Db 4460 TCACGTGATGTCCTCTTTGATTCAGAGAGATTACCAAAAAGTGTAACATTGTAAGG 4519
Qy 1826 atttccaatctgcgcagagaaatattcaataatggaagtgactgttagaagtg 1885
Db 4520 ATTTTCCAAATTCGCCAGGAGAAATATTCAAATTAATGACAGTACGTAAAGATG 4579
Qy 1886 ggcacaataatcaagatccctcgctgctgacccgacttaacttaagttagttagtg 1945
Db 4580 GGGCAATTAATTCAGATTCCTGCTGCTGACCCGCTATTAATCTAGTTTGTTAATAGG 4639
Qy 1946 agagagatctagcttcagagactcattggccctcctccatctgcttaaaaaagactgtag 2005
Db 4640 AGAGAGATCTAGCTTCAGAGACTGTTGGCCCTCTCCATCTGCTACAAAGAACTGTAG 4699
Qy 2006 atcaagaagagaacacagataatgtcagaagaagaaatgcaatccgttcttctgtattg 2065
Db 4700 ATCAAAAGAGAAACCATATATGTCAACAAAGAGAAATGTCATCTTTCTTATTTG 4759
Qy 2066 atgagaacccgaagctgtgtactcacaagaatatataacgcttctcccaatccagtg 2125
Db 4760 ATGGAACCCGAAGCTGTACTCTCAGAGAAATATACAAAGCTTTCTCCCAATCCAGCTG 4819
Qy 2126 gaagtgcagcttgaagatccagaggttccaaagcctccaaacatgcaagcatatgct 2185
Db 4820 GAGTGCAGCTTGAAGATCCAGAGTCCAAAGCTCCAAACATGCAAGCATCAATGCT 4879
Qy 2186 atgtttttagtagtttgcagttgtcagttgttgcagtgagtgagtaactgtatctc 2245
Db 4880 ATGTTTTTGTAGTTTGCACATTTGTCATTTGTTGTCATGTAGGTGGCATCTGCTGATTC 4939
Qy 2246 taagcatltgagcagacagagactcctctgtcctcctcctcgtgatatcccttaaac 2305
Db 4940 TAAACATTTGAGACACAGACTGACTCTTCTGTGCTTCTTCTCTGGATATACCTTCAAC 4999
Qy 2306 acaaaatggtctatgaaagacacatcccatatcccatctcagagaaactgtctca 2365
Db 5000 ACAAAATGCTCTATGAAGACACTACCCATTTCCCATTTCTCAGAGAAAGTGTCTTCA 5059
Qy 2366 tgtcgtatggaaaacccaggtctatgatatctggggtgcacaactcagacttcgagaa 2425
Db 5060 TGTGATGAGAAAACCCAGGCTCTATGAGATCTGGGGTGCACAACTCAGACTTTCGGAACA 5119
Qy 2426 gaagcatgacgccttactgaaggttctagttgtgacaagaacacatggtgattattag 2485
Db 5120 GAGGCAATGACCGCTTACTGAAGTTTCTAGTTGTGACAAAGAACACTGGGATTAATTAAG 5179
Qy 2486 aggaacgtltagaatattcagcatactgtcgaataaaacatgcatgaaacaa 2545
|||||
```

---

```
Db 5180 AGACAGTTATGAAGATATTTACAGCATACTTGTGAGTAAAAACATGCCATGACCAA 5239
Qy 2546 gaagcttctccagaatccaagaacccctagcactaggcaaaagcaattatgccaacc 2605
Db 5240 GAAGCTTCTCCAGAAATTCAGAGACACCTTAGCATTAGCAAAAACAAATTAATGCCACC 5299
Qy 2606 caccagcttgaaaagccatcaaggggaataactgttactactcttcagccaatcaag 2665
Db 5300 CACCATGTTGAAACGCCATCAAGGGAATTAACCTGTAATCTTCAATGATCAATCAAG 5359
Qy 2666 aggaaatltagctatgatalcaccatcaagttgaatgaagaagaagatttgaactt 2725
Db 5360 AGAAATTTGATATGATGATATCATATCAGTTGAATGAATGAAGAAAGATTTGACATTT 5419
Qy 2726 atgataagatgaaatcaagagcccgcaagcttccaagaagaacacagacataattta 2785
Db 5420 ATGATGAGAGTAAATCAGAGCCCGCACCTTTCAAAAGAAACAGACACTATTTTA 5479
Qy 2786 ttgctgcaatggaagagctcctgggattatgagatgagtgagtcctcccaatgtltaaga 2845
Db 5480 TTGCTGCAATGAGAGAGCTCTGGGATTAATGGATGAGTACTCCCAATGTTTAAGAA 5539
Qy 2846 acagagctcagagtgagcagtgctccctcagttcaagaagaagttgtttccaagaaattactg 2905
Db 5540 ACAGGCGTCAGAGTGGCAGTGTCCCTCAGTTCAAGAAAGTTGTTCCAGGAATTTACTG 5599
Qy 2906 atggctccttactcagcccttataccgttggagaactaaatgaacatttggactcctg 2965
Db 5600 ATGGCTCTTACTACACGCCCTTAATACCGTGAAGCACTAAATGAACATTTGGGACTCCTGG 5659
Qy 2966 ggcacatataagaagcaagttgaagataatcatggtgaacttcaagaataagcct 3025
Db 5660 GGCCATTAATAGAGCAGCAAGATTGAAGTAATATCATAGTACTTTTGAATATAGGCT 5719
Qy 3026 ctgctccctatcccttactatctagccttatttctttagaagaatcagagcaagag 3085
Db 5720 CTGCTCCCTATTCCTTCTATCTAGCTTATTTCTTAATGGAAGATCAAGGCAAGGAG 5779
Qy 3086 cagaacctagaanaaacttgtcagaagccttaatgaacccaaacttactttagaagtgc 3145
Db 5780 CAGAACCTAGAAAAAATTTGTCAAGCCTAATGAACCAAAACCTTATTTGGAAAGTGC 5839
Qy 3146 aacatctatggcaccacaaagaatagtttgaacttgcacaaagccttggcttattctctg 3205
Db 5840 AACATCTAATGGCACCACCTAATAGATGATTTGACTGCAAAAGCTGGGCTTATTTCTTG 5899
Qy 3206 atgttgacctggaanaagatgtgactcagcagcctgcatltagacccctctgtctgcaca 3265
Db 5900 ATGTTGACCTGGAATAAGATGTGACTCAAGCCTGATTTGGACCCCTCTGGTCCGACA 5959
Qy 3266 ctaaacactggaacctgtctcaatggagaacaagtgaacagtaacagaaattgtctgtttc 3325
Db 5960 CTAAACACACTGAACCCGTGCTATGGAGACAAAGTGAACAGAAATTTGCTCTGTTTT 6019
Qy 3326 tcaaccatctttagtgaagacaaaagctgtgtactcactctgtaaaaatgaaagaanaagcga 3385
Db 6020 TCACACATCTTTGATGAAACCAAAAGCTGTACTTCACTGTAATAATATGAAAGAAACTGCA 6079
Qy 3386 gggctccctgcaatalccagaatggaagatcccaactttaaagagaattatcgtctccatg 3445
Db 6080 GGGCTCCCTGCATATATCAGATGGAAGATCCCACTTTTAAAGAAATTAATTCGTTCCATG 6139
Qy 3446 caatcaatggtctacataatgataactactcgttgtagtaatggtcagagatcaaaaga 3505
Db 6140 CAATCAATGCTCTACATATATGATGATACACTACCTGGCTTAGTAATGGCTCAGAGATCAAAAGGA 6199
Qy 3506 ttgatgtatctgtcgaagatggcagcaatgaanaacatcaatctcatctattttagtg 3565
Db 6200 TTGAGTGTATTCCTCAGATGGGACAGCAATGAAACATCCATTTATTCATTTTCAAGTG 6299
Qy 3566 gacatgttctcactgtltagaaaaaagagagtlataaaatggaactgttaaatctcctatc 3625
Db 6260 GACATGTTTCACTGTACGAAATAAGAGAGATTAATAATGGCACTGTACAAATCTCATATC 6319
|||||
```



QY 3626 caaggctgtttcttgagacagctggaaatgltacacatcccaagcttgaaatttgccggttggaat 3685  
|||||  
Db 6320 CAGGTGTTTTGAGACGATGGAAATGTTACCATCAAAAGCTGGAAATTGGCGGGTGAAT 6379  
QY 3686 gacctattgacgagacatcacatgctggatgagcacactttctggtgtacagcaata 3745  
|||||  
Db 6380 GCGTTATTTGGGAGACATCTACATGCTGGAGTAGACACACTTTTCTGCTGATCACCAATA 6439  
QY 3746 agtgcagaccctctgggaatgctcttgacacatcagaagatttcaagattcaagactt 3805  
|||||  
Db 6440 AGTTCAGACTCCCTGGGAATGGCTTCTGCACACATTAGAGATTTTCAGATTACAGCTT 6499  
QY 3806 caaggacaatatgagacagtgggcccccaaaagctggccagacattcatatctccgagatca 3865  
|||||  
Db 6500 CAGSACATATGAGACAGTGGGCCCCCAAAAGCTGGCCAGACTTCATTATTCGGATCAATCA 6559  
QY 3866 atgccttgagacacacaggaagcccttctcttgatcacaagtgatctgttgccacaatga 3925  
|||||  
Db 6560 ATGCTTGAGGACCAAGAGGCCCTTTCTTGGATCAAGGTGATCTGTTGGCACCAATGA 6619  
QY 3926 ttattcaagcatcaagaaccccaaggctgcgcgtcgaagattctccagccctcactctctc 3985  
|||||  
Db 6620 TTATTCACGGCATCAAGACCCAGGGTGGCCCTCAGAACTTCTCCAGCCTTCAATCTCTC 6679  
QY 3986 agttatcatcatglatagctcttgatggaaagtgccagactatcgaagaaattcca 4045  
|||||  
Db 6680 AGTTATCATCATGTATAGTCTTTGATGGGAAGAAAGTGCGACAGCATATGCGAGAAATTCGA 6739  
QY 4046 ctggaaccttaatgctctctcttggaatgtggaatcactcggggaataaacaatattt 4105  
|||||  
Db 6740 CTGGAACCTTAATGATGCTCTTTTGGCAATGTGATTCATCTGGGATTAACCAATATT 6799  
QY 4106 ttacccttccaatattgctgatacatccgtctgaccacacatataagcatctgca 4165  
|||||  
Db 6800 TTAACTTCATATTATGCTGTGATACATCCGTTTGACCCAACTATTAATACATTTCCGA 6859  
QY 4166 gcaactctgcatagagttgaatggcctgctgattcaatagttgagcaatgacatgggaa 4225  
|||||  
Db 6860 GCACTCTTCGCGATGAGTGTGATGGGCTGTGATTTAAATAGTTGACGATGCAATTTGGGA 6919  
QY 4226 tggagaaataagaacataatcagatgacagattactgctcactcacttaccataatgt 4285  
|||||  
Db 6920 TGGGAGATTAAGCAATATCAATGATGACAGATTACTGCTTCACTTCCATTTACCAATAGT 6979  
QY 4286 ttgcacctgtgctccttcaaaaagctgactcactcactcacaaggaggaagtatgctgga 4345  
|||||  
Db 6980 TTGCGACCTGGTCTCTTCAAAAGCTGCACCTTCACCTCCAAAGGAGGATATGCTTGA 7039  
QY 4346 gaacctgagtgaataatccaaaagagtgctgcaagtggaacttccagaagaatgaaag 4405  
|||||  
Db 7040 GACCTCAGGTGAATATCAAAAGAGTGGCTGCAAGTGGACTTCCAGAAAGCAATGAAG 7099  
QY 4406 tccagagaagtaactactcagaggagtaaaattctgcttcaacagcaatgatatgaaagggt 4465  
|||||  
Db 7100 TCAAGAGGATACATCTCAGGAGTAAATCTCTGCTTACACAGCAATGATGGAAGAGT 7159  
QY 4466 tcccatcctcagcagcagcaaatgccaatcagtggaactccttcttcagaatgcaag 4525  
|||||  
Db 7160 TCTCATCTCCAGCAGTCAAGATGGCATTCAGTGACCTCTTTTTCAGAAATGGCAAG 7219  
QY 4526 taaagtttccagggaatcaaaactcctcacaacctggtggaactcctcagaaccac 4585  
|||||  
Db 7220 TAAAGGTTTTTCAGGAAATCAAGACTCCTTCACACCTGTGTGAATCTCTAGACCCAC 7279  
QY 4586 cgttactgactcgtactcctcgaattcaacccccagagttgggtgacacagatgtccctga 4645  
|||||  
Db 7280 CGTTACTGACTCGCTACTCTTGAATTCACCCCAAGTGGGTGACCAAGATTTGCCCTGA 7339  
QY 4646 ggtatgaggtctcgggtcggagacagagacactactagactcgaaggatctc 4699  
|||||  
Db 7340 GGATGGAGGTTCTGGGCTGCGAGGACAGGACCTTACTGAGGGTGGCCACTGC 7393

RESULT 2  
AR110040 4670 bp DNA PAT 14-FEB-2001  
LOCUS AR110040  
DEFINITION Sequence 41 from patent US 6114148.  
ACCESSION AR110040  
VERSION AR110040.1 GI:12826316  
KEYWORDS  
SOURCE  
ORGANISM Unknown.  
REFERENCE  
AUTHORS 1 (bases 1 to 4670)  
TITLE Seed, B. and Haas, J.  
JOURNAL High level expression of proteins  
FEATURES  
SOURCE Patent: US 6114148-A 41 05-SEP-2000;  
1. 4670  
Location/Qualifiers  
BASE COUNT 1325 a 1072 c 1036 g 1237 t  
ORIGIN  
Query Match 91.5% Score 4357.6; DB 9; Length 4670;  
Best Local Similarity 99.8% Pred. No. 0;  
Matches 4363; Conservative 0; Mismatches 9; Indels 0; Gaps 0;  
QY 328 gccaccagaagaactaccctgggtgacgtggaactgcatgtggaactatggaactatgcaagtgat 387  
Db 84 GCCACGACAGATACTACTCTGGGTGACAGTGGATGGAAGCTGATGGGACTATATGCAAAAGTAT 143  
QY 388 ctggtgagctgctgtggaagcaaatctccctcagatgagcaaaatcttccattc 447  
Db 144 CTCGGTGAAGTCCCTGTGAGCAGCAAGATTTCCTCAGAGCCAAATCTTTTCATTC 203  
QY 448 aacacactaagctcgtgtacaaaaagactctgtttgtagaatcgaagttacacttccaac 507  
Db 204 AACACCTTAGTCTGTATCAAAAAAGACTCTGTTTGAATTCACGATACACCTTTTCAAC 263  
QY 508 atgcgttaagcaagaccacccctgagtggtctgctgaagctccacacacagagctgaagtc 567  
Db 264 ATCGCTAAGCCAGGCGCCACCTGGATGGGTGCTGCTAGTCTTACCATCCAGGCTGAGGTT 323  
QY 568 tatgataagtgatcattacacttaagaacatggtctccactcgtcagtcattcagtc 627  
Db 324 TATGATACAGTGCATATACACTTAAGAACATGGCTTCCCTGACAGCTTCATGCT 383  
QY 628 gttagtatactcctctgaaaagctctctgaggaactgaataatgataatgataagaccagca 687  
Db 384 GTTGGTGTATCTTACTGAAAGCTTCTGAGGAGCTGAAATGTGATGATCAAGACAGTCAA 443  
QY 688 agggagaagaagaatgataaagctctccctgtgtgaagccatataatgctgcaagtc 747  
Db 444 AGGGAAGAAAGATGATTAAGTCTTCCCTGTGGAGCCATATATATGCTGGCAGTTC 503  
QY 748 ctgaaagaagaatggtccaaatgagctctgaaacccactggtccttactcactaataatctctc 807  
Db 504 CTGAAGAAGAAATGTGTCATGAGCTGACCCCTGAGCCCTTACTACTCATATATCTTCT 563  
QY 808 catgtgagccctgtaaaagacttgaatcagccctcactcactcactcactcactcactcactc 867  
Db 564 CATGTGACCTGGTAAAGACTTGAAATTCAGGCTCATGTGAGGCTTACTAGATGATGTA 623  
QY 868 gaaggagactctgcccagaagaagaacacagacccctgacaaattataactactcttgc 927  
Db 624 GAAGGAGTCTGGCCCAAGGAAAGACACAGACCTTGCAAAATTTATATACATCTTTTGTCT 683  
QY 928 gtatttgaatgaaggaaagtctggcactcagaacaaagaactccttgatgcaagatag 987  
Db 684 GTATTGTATGAAGGAAAGTTGGCAGCTCAGAAACAAAGAACTCTGTATGACAGATAGG 743  
QY 988 gatgctgatactgctcgggctcgtgctcctaaatgacacagatcgaatggtatgtaaagag 1047  
Db 744 GATGCTGATCTGTGCTGGGCTTGGCTTAAATGCAACAGTCAATGATGATTAACAGG 803



QY 1048 tctctccaggtcttgattggaatgcacaggaatcaatctatctggatctgattggaatg 1107  
|||||  
Db 804 TCTCTGCCAGGTCGTGATTGGATGCCACAGGAATCAGCTCTTTGGCATGTGATTGGATG 863  
QY 1108 ggcaccactcttgaaagtgactcaataltccctcgaaggctcacacatctctgttgaggaa 1167  
|||||  
Db 864 GGCACACTCTCTGAAGTGCATCAATATCTCGAAGGTCAACATTTCTTGAGGAAAC 923  
QY 1168 catgcgcagggctctcttggaataatctgcgaataactctctactctgtctcaaacctcttg 1227  
|||||  
Db 924 CATGCCAGGGGCTCTGGAAATCTCGCAATTAACCTTCTACTGCTCAAAACCTCTTG 983  
QY 1228 atggaccttgacagattctactctgtttgtcatatctctcccccacacatgagagatg 1287  
|||||  
Db 984 ATGGACCTTGACAGATTCTCTCTGTTTGTATATCTCTTCCCAACATGATGGCATG 1043  
QY 1288 gaagcttatgtcaaaagtagaagctgtccagaaggaaaccccaactacgaatgaaaaataat 1347  
|||||  
Db 1044 GAAGCTTATGTCAAAAGTAGACAGCTGTCCAGAGAACCCCAACTACGAATGAAAAATAT 1103  
QY 1348 gaagaagcgggaagactatgatgatctctacgtatcttgaaatgagatggtgtaagttc 1407  
|||||  
Db 1104 GAAGAAGCGGAAGACTATGATGATGATCTTACTGATTCTGAAATGAGTGTGTCAGGTTT 1163  
QY 1408 gatgataacaactctctctctcttataccaaattgcgcagtgtgcaagaagaacatccctaa 1467  
|||||  
Db 1164 GATGATGACACTCTCTCTCTCTTATCCAAATGTGCTGAGTGCAGGAAGAACATCCCTPAA 1223  
QY 1468 acttggtgtacatlatcatctgtcgtcgaaggaggagactggtgactatgtctccctatgtctc 1527  
|||||  
Db 1224 ACTTGGGTACATTAATTTGGTGTGAAGAGAGACTGGGACTATGCTCCCTTAGTCTC 1283  
QY 1528 gcccccagatgacagaagttataaaagtcaataattgaaacaatggccctcagcggatctggt 1587  
|||||  
Db 1284 GCCCCCATGACAGAAAGTTATAAAGTCAAATTTGAAACAATGGCCCTCAGCGGATTGGT 1343  
QY 1588 aggaagtaacaaaagtcgcgaatttatggacatacacagatgaacactttaagaccctgaa 1647  
|||||  
Db 1344 AGGAAGTACAAAAGATCCGATTTATGGCATACACAGTGAACCTTTAAGACTCGTAA 1403  
QY 1648 gctatcaagcatgaatacagaatcttggaacttacttataatgggaagtgtggaacaa 1707  
|||||  
Db 1404 GCTATTGACGATGAATCAGGAATCTTGGACCTTACTTATGSGGAAGTTGAGACACA 1463  
QY 1708 ctgtgtattatatttaagaatacaagcaagcagacacataacatctacccctcaaggaatc 1767  
|||||  
Db 1464 CTGTTGATTATATTAAAGAAATCAAGCAAGCAGACCATATAACATCTACCCACGGAATC 1523  
QY 1768 actgattccgtctcttgatctcaagagatctaacaaaggtgtaaaacatttgaagat 1827  
|||||  
Db 1524 ACTGATGTCCCTCTTGTATTTCAAGAGATTAACCAAAAGGTGTAACATTTTGAAGAT 1583  
QY 1828 ttccaactctgcagaggaataatltcaaatataatgagcagtgactgtagaagatggtg 1887  
|||||  
Db 1584 TTTCCAATTTGCCAGAGGAATATCAAAATATAATGAGACAGTGAAGATGGG 1643  
QY 1888 ccaactaatctagatctctgtgtgcgtgaacccgcctatacctcctagtctcgtaataatgag 1947  
|||||  
Db 1644 CCAACTAAATCAGATCTCGGTGCTTACCGCTATTAACCTAGTTTCGTTAAATATGAG 1703  
QY 1948 agagatctagctcaggaactatgcccctctctcatctgcgtacaagaagatcgttagat 2007  
|||||  
Db 1704 AGGATATAGCTTACAGACATCATTTGGCCCTCTCTCATCTGTCTCAAAAGATCTGTAGAT 1763  
QY 2008 caaagaggaacacagataatgtcagacaagaggaatgtlcalcctgttcttctgtatgtat 2067  
|||||  
Db 1764 CAAAGAGGAACACGATTAATGTACAGACAAGAGAAATGTCAATCTGTTTCTGTATTTGAT 1823  
QY 2068 gagaacccgaagctgtgtacccctcacagagaataataaagcttctccccaatccagctgga 2127  
|||||  
Db 1824 GAGAACCGAAGCTGTGACTCTACAGAGAAATATACAAACGCTTTCTCCCAATCCAGCTGGA 1883

QY 2128 gtcagcttgaggaatccagagttccaaagcctccaacatcatatgcacagcatcaatggtcat 2187  
|||||  
Db 1884 GTGACGCTTGAGGATCCAGAGATTCCAGAGCTCCACATCATGACATCAATGCGTAT 1943  
QY 2188 gttttgatagtttgcaagtggtcaggttctgtttgtatgaaagtgatgactgtatattcta 2247  
|||||  
Db 1944 GTTTTGTATAGTTTGCAAGTTGTGAGTTGTTGTGATGAGGTGGCATACTGTGTATTTCTA 2003  
QY 2248 agcatlgagacagagactgagacttctctctgtctctctctctctctgatataccctcaaac 2307  
|||||  
Db 2004 AGCATTTGAGACAGACACTGACTTCTTCTTCTTCTTCTTCTGATATTAACCTTCAAAAC 2063  
QY 2308 aaaaatgtctatgagaacacactcaacctatcccatctcccaagagaagactggtcttatg 2367  
|||||  
Db 2064 AAAATGCTATGAGACACACTACCTATTCCTCATTTTCAGAGAAACTGTCTTATG 2123  
QY 2368 tcgaatgaaaaacccaggtctatagatctctggtgtgcacaactcagaactcttcggaacaga 2427  
|||||  
Db 2124 TCGATGGAAAAACCCAGGCTATGATGATCTGGGGTGCCCAAACTCAGACTTTCGGAACAGA 2183  
QY 2428 ggcatagcgcgtctactcgaaggttctcagtgtgtgaacaagaactggttgattatcagag 2487  
|||||  
Db 2184 GGCATGACCCGCTTACHTGAAGGTTTCTAGTTGTGACAAAGAACACTGGTATTAATTACGAG 2243  
QY 2488 gacagttatgaagatatttcaagcaactctgtgtggttaaanaaaatgscattgaaaccaaga 2547  
|||||  
Db 2244 GACAGTTATGAGATATTTTACGACTTCTGCTGATGATAAACAATGACATTTGAACACAGA 2303  
QY 2548 agcttctccagaatltcaagaacacccctagacacttgagaaagaacaaatlaatgcccacca 2607  
|||||  
Db 2304 AGCTTCTCCAGAAATTCAGACACACCTTAGCACTAGGCAAAAGCAATTAATGCGACCCCA 2363  
QY 2608 ccagctcttgaaagcgcatacgaaggaataactcgtactactctctcagtaagaacagag 2667  
|||||  
Db 2364 CCACTCTTGAACCCCATCAACGGGAATTAACCTGCTACTCTCTCAGTCAAGTCAAGAG 2423  
QY 2668 gaattgacctatgataatcacataatcagttgaaatgaaagaaggaatttgacattat 2727  
|||||  
Db 2424 GAAATTTACTATGATGATATACCATATACGTTGAAATGAAAGAAAGATTTTGGACATTTAT 2483  
QY 2728 gatgaagatgaaaaatcagaagccccgcagctltcaaaagaanaacagacacatattat 2787  
|||||  
Db 2484 GATGAGGATGAATAATCAGAGCCCCCGCAGCTTCAAAAAGAAACACGACACTATTTAT 2543  
QY 2788 gctgcagatgagagggctctggtatataatgagatgagatgctcccaactgttctaaagaac 2847  
|||||  
Db 2544 GCTGCAGTGAAGAGGCTCTGGGATTATGAGATGAGTGTGCTCCACATGTTCTAAAGAAC 2603  
QY 2848 agggctcagaatgagctgtccctcagttcaagaagaattgttccaggaattactgat 2907  
|||||  
Db 2604 AGGGCTCAGAGTGGCAGTGTCCCTCACTTCAAGAAAGTTGTTCAGGAATTTACTGAT 2663  
QY 2908 ggtctcttactcagcccctataacggtggaactaaatgaacatttgggaactcttggtg 2967  
|||||  
Db 2664 GGCTCTTTACTCAGCCCTTATACCGTGAGAACTAAATGAACATTTGGAGCTCTCGGG 2723  
QY 2968 ccatataatagagcagaagtttgaagataataatcatgttgaacttctcaagaatcagggctct 3027  
|||||  
Db 2724 CCATATATTAAGAGAGAAAGTTGAAGATTAATCATGTGTAACCTTCAAAAATCAGGCTCT 2783  
QY 3028 cgtccctactctctatcttaagcctattcttaatgaagaagaatcagaagcagaagaga 3087  
|||||  
Db 2784 CGTCCCTATTTCTTATTTAGCTTATTTCTTATTAAGAGAAATCAGAGCAAGAGGACA 2843  
QY 3088 gaaactagaanaaaactltgtcaagcctaaatgaacaaacaaacttacttltggaagtgcga 3147  
|||||  
Db 2844 GAACCTGAAAAAACTTTGTCAAGCCTAATGAACCAAAACTTACTTTGGAAGTGCA 2903  
QY 3148 calcataatgcaacccataaagaatgagttgactgtgaagacccgtggcttatctctgat 3207  
|||||  
Db 2904 CATCATATGGACCCCACTAAAGATGAATGATGTGCAAAACCTGGGGCTTATTTCTCTGAT 2963  
QY 3208 gttgacctggaanaagatgtgcactcagcgctgtatgtgaacccctctgtctgcacact 3267



|||||  
Db 2964 GTTACCGGAAAAAGATGTGCACACGAGCCGATGAGCCCTCTGTGTCACACACT 3023  
Oy 3268 aacacactgaacccctgctcatcagggagacaaatgacagtaacagaaattgctctgttttc 3327  
Db 3024 AACACACTGAACCCCTGCTCATGCGGAGACAGTACAGGAATTTGCTCTGTTTTTC 3083  
Oy 3328 accactcttga tgaagcaaaagctgtaactgaataata tgaagaaactgcag 3387  
Db 3084 ACCATCTTGTATGAGACCAAAAGCTGTACTGATAAATATGAAAGAACTGCAGG 3143  
Oy 3388 gctccctgcaatatccagatggaagatcccaactttaagagaatatacgttccatgca 3447  
Db 3144 GCTCCCTGCAATATCCAGATGGAATCCACTTTTAAAGAAATATTCGCTTCATGCA 3203  
Oy 3448 atcaatgctcaataa tgaataactaactgctgctgaatgaatgctcaaggaatgaat 3507  
Db 3204 ATCAATGGCTACATATATGATACACTACCTGGCTTATGATGGCTGAGATCAAGGAT 3263  
Oy 3508 cgaatgatactgctcagcatgagcagcaatgaanaacatccatctatccatcagtgga 3567  
Db 3264 CGATGATATCTGCTCAGCATGGGAGCAATGAANAACATCCATTTATTCATTTCACTGA 3323  
Oy 3568 caatgctcactgtaagaaaaaagaagagataaaatggaactgtaacaaatccatcca 3627  
Db 3324 CATGTGTTCACTGTACGAAAAAAGAGATATAAATGCGACTACATCTTATCA 3383  
Oy 3628 gggtttttgagacagtgaaatgttaccatccaaagctggaattgacgggtggaatgc 3687  
Db 3384 GGTTTTTTGGAGACAGTGGAAATGTTACCATCCAAAGCTGGAAATTTGGCGGTGGAAATGC 3443  
Oy 3688 ctatctgagcagcatctcaatgctgga tgaagacacttctctgtgtgtaacagcaataag 3747  
Db 3444 CTATATGGCGAGCATCTCATGTGCGATGAGCACACTTTTCTGGTGTACAGCAATAAG 3503  
Oy 3748 tgttaagactccccggggaatgcttcttgacacatlaagatcttccagatlaacacttca 3807  
Db 3504 TGTGAGACTCCCTGGGAAATGGCTTGTGACACATTAAGATTTTCAATTCACCTTCA 3563  
Oy 3808 ggaacataatgacagtgagccccaagcttgccagactctatctccagatcaatcaat 3867  
Db 3564 GGAACAATATGACAGTGGGCCCCAAAGCTGGCGACACTTCATTTCCGAGCAATCAAT 3623  
Oy 3868 gactgagacacaaagagcccttctctgga tcaagtgatgacgtgtgacacaaatgalt 3927  
Db 3624 GCCTGAGACACCAAGAGACCCTTTCTTGATCAAGTGTGATCTGTGGACCAATGATT 3683  
Oy 3928 attacaggaatcaagaacccaaggtgtgcccgtcaagaagttctccagctctacatctcaag 3987  
Db 3684 ATTCACGCAATCAAGACCCAGGTCGCCGTCAAGAGTTCTCCAGCCTCTACATCTCTCAG 3743  
Oy 3988 tttaacatcatgatagtctgataatggaagaatgagacactatccgaagaaatccact 4047  
Db 3744 TTTATCATCATGATATAGCTTGTGATGGAAAGATGGCAGACTTATTCGAGAAATTCACAT 3803  
Oy 4048 ggaacctaataatgctctctcttgcaaatg tgaatcaatctggaataaacaataatctt 4107  
Db 3804 GGAACCTTAATGTCTCTTGTGGCAATGTGATCATCTGGATTAACACAAATATTTT 3863  
Oy 4108 aacacctcaatatctgctcgataatccgcttgcaacccaactcatatagacatctgcagc 4167  
Db 3864 AACCTCAATTAATGTCTGTGATCATCTGTCGTTGACCCAACTATTATAGCTTTCGACGC 3923  
Oy 4168 accttcgaatgagatgagctgagctgataatgaatgctgaacatgcaatctggaatg 4227  
Db 3924 ACTTTCGCAATGAGATTCATGGCTGTGATTTAAATAGTTGAGAGATGCATTCGGGAATG 3983  
Oy 4228 gagaatgaagaatatacagatgacagata tcaatgcttcaatccatcttacaataatgltc 4287  
Db 3984 GAGAGTAAAGCAATATCATGATGACAGATTAATCTGCTCATCTTACCAATATGTTT 4043  
Oy 4288 gcaacactggtctccttcaaaagctcgacttcaactccaaggaaggaatgctccggaga 4347  
|||||

Db 4044 GCCACCTGGTCTCTTCAAAAAGCTGCACCTTCACAGGAGAGTAATGCTGGAGA 4103  
Oy 4348 cctcagatgaataatccaaaagatgctgcaagtgcagctccagaagaatgaagtc 4407  
Db 4104 CCTCAGGTGAATATATCCAAAAGATGGCTGCAGATGCACTTCAGAAACATGAATGTC 4163  
Oy 4408 acagaagaataactcaaggaaglaaataatctcgtcttaccagatglatggaagatlc 4467  
Db 4164 ACAGAGTACTACTACAGGAGTAATAATCTGCTTACACAGATGATGTGAAGAGATTTC 4223  
Oy 4468 ctcatctccagagatcaagaatgagccatgaatgagctctcttttccgaatggaagta 4527  
Db 4224 CTTCATCTCCAGAGTCAAGATGCGCATCATGATGACTCTTTTTCAGAAAGCAAGTA 4283  
Oy 4528 aaggttttcaaggaataatcaaacactctcaacactggtggaactctcagaccacg 4587  
Db 4284 AAGGTTTTTCAAGGAATATCAAGACTCTTCACACTGTGGTGAACCTCTTAGACCCACG 4343  
Oy 4588 ttaactgactcgtactacatcttcaatccaccagaagttggtgacacagatgcccctgag 4647  
Db 4344 TTACTGACTCGCTACTCTTCGAATTCACCCCGAGATTGGTGACACAGATTGCCCTGAGG 4403  
Oy 4648 atggaagttctggtgctgagagcaagaagactctactgactcagagagatlc 4699  
Db 4404 ATGAGAGTTCTGGCTGGAGCACAGGACCTCTACTGAGGGTGGCCACTGC 4455  
  
RESULT 3  
108644  
LOCUS 108644 4278 bp PAT 02-DEC-1994  
DEFINITION Sequence 4 from Patent WO 8800831.  
ACCESSION 108644  
VERSION 108644.1 GI:588650  
KEYWORDS .  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 4278)  
AUTHORS Pasek,M.P.  
TITLE DNA SEQUENCES CODING FOR MODIFIED FACTOR VIII:C AND MODIFIED FACTOR VIII:C-LIKE POLYPEPTIDES AND PROCESSES FOR PRODUCING THESE  
POLYPEPTIDES IN HIGH YIELDS  
JOURNAL Patent: WO 8800831-A 4 11-FEB-1988;  
FEATURES  
source 1..4278  
BASE COUNT 1244 a 942 c 946 g 1146 t  
ORIGIN  
  
Query Match 87.1%; Score 4148.8; DB 10; Length 4278;  
Best Local Similarity 97.9%; Pred. No. 0;  
Matches 4270; Conservative 0; Mismatches 2; Indels 90; Gaps 3;  
  
Oy 328 gccacagaagaatctactcctggtgagtggaactgtaatggaactatataatgaaatgat 387  
Db 1 GCCACCAAGAAATACTACTGCTGGTGCAGTGAACGTGCATGGGACTATATGCAAAAGTAT 60  
Oy 388 ctgggtgagctgcccgtgtgagcagaagatttctctcgtagagtgagcaaaatctttccatc 447  
Db 61 CTCGGGTGAGCTGCCGTGTGACGCAAGATTTCTCTTAAGTGCMAAATCTTTTCCATTC 120  
Oy 448 aacacctcagtcgtgtacaaaagactcgtgttgaatcaacaggttcaactcttcaac 507  
Db 121 AACACTCAGTCTGTATCAAAAAGACTCTGTTGTAGAAATTCACGATCACTTTTCAC 180  
Oy 508 atcgttaagcaagagcaacccctgagatggctgctagagttctaccatcaagctgaggt 567  
Db 181 ATCGCTAAGCCAAAGGCCCTGATGGGTCTGTAGTCTTACCATTCAGGCTGAGGTT 240  
Oy 568 tatgataagtgatcatcaacttaagaaatgcttccatccatcgtcagatcttcatgct 627  
Db 241 TATGATACAGTGTCTATTAACCTTAAGAACATGGCTTCCATCTGTACGTTCATGCT 300  
|||||



QY	628	gttgggtatccctacttggaaagcttcttgaggagctgaaatacgtatgcacgaccgaa	687
Db	301	gttggtgtatcttaccatgcgaagaaagcttctgaggagctgaaatacgtatgcacgaa	360
QY	688	agggagaagaagatgataaagtccttcctcgtgtgaaagcacaatacgtctcgtcagtc	747
Db	361	agggagaagaagaatgataaagtccttcctcgtgtgaaagcacaatacgtctcgtcagtc	420
QY	748	ctgaaagaagaatgtccaaatgagcctctggcccacgtctgcttaccatcatatccttc	807
Db	421	ctgaaagaagaatgtccaaatgagcctctggcccacgtctgcttaccatcatatccttc	480
QY	808	catgtggaccctggtaaaagacttgaatctaaagcctaatgtgagccctactaglatgata	867
Db	481	catgtggaccctggtaaaagacttgaatctaaagcctaatgtgagccctactaglatgata	540
QY	868	gaagaggagctctggccaaaggaaaaagacacagaccttgcacaaattatctactcttgc	927
Db	541	gaagaggagctctggccaaaggaaaaagacacagaccttgcacaaattatctactcttgc	600
QY	928	gtattcgtatgaaggaaagattgagcactcagaagaacaaagacctccttbatcagatag	987
Db	601	gtattcgtatgaaggaaagattgagcactcagaagaacaaagacctccttbatcagatag	660
QY	988	gatgctgcacatcgtctcggagcctggccataaatgacacagtaacgtatgtatgataacag	1047
Db	661	gatgctgcacatcgtctcggagcctggccataaatgacacagtaacgtatgtatgataacag	720
QY	1048	tcctctg----ccaagtcctgatttgaatgcacaaagaatacagtcatactgtatgattgata	1104
Db	721	tcctctgcttaccagagctctgatttgaatgcacaaagaatacagtcatactgtatgattgata	780
QY	1105	atgggacacacactcctctgaagtgacactcaataatctccctcogaagtgacacacttctcgtgag	1164
Db	781	atgggacacacactcctctgaagtgacactcaataatctccctcogaagtgacacacttctcgtgag	840
QY	1165	aaccatctgcgaagcgctccctcttgaaatctctgcgaataactcttccttactgctccaacactc	1224
Db	841	aaccatctgcgaagcgctccctcttgaaatctctgcgaataactcttccttactgctccaacactc	900
QY	1225	ttgatggaccttggacagatttctacagtttcttgcatactctctccacacaatgataatggc	1284
Db	901	ttgatggaccttggacagatttctacagtttcttgcatactctctccacacaatgataatggc	960
QY	1285	atggaagcttatgtlcaaaagtacagcctgtlccagaagaaaccccaactacgaatgaaaaat	1344
Db	961	atggaagcttatgtlcaaaagtacagcctgtlccagaagaaaccccaactacgaatgaaaaat	1020
QY	1345	aatgaagaacggaaagacataatgataatgataatcttaccagatcttcgaataatgataatgataatg	1404
Db	1021	aatgaagaacggaaagacataatgataatgataatcttaccagatcttcgaataatgataatgataatg	1080
QY	1405	tttgatgatgacacactctcccttcccttccataccaaatctgcctcagtttgcgaagaagaatcatct	1464
Db	1081	tttgatgatgacacactctcccttcccttccataccaaatctgcctcagtttgcgaagaagaatcatct	1140
QY	1465	aaaaacttggttacattacatctgctctgtcgtgaagaggagaccttggactatgctcccttaagtc	1524
Db	1141	aaaaacttggttgatattatcatttgcctggaagaggagaccttggactatgctcccttaagtc	1200
QY	1525	ctgcgccccgcatcacagaagattaaagaatcaatttgaacaaatggcgccctcagcgagatt	1584
Db	1201	ctgcgccccgcatcacagaagattaaagaatcaatttgaacaaatggcgccctcagcgagatt	1260
QY	1585	gttatagaagtacaaaaaagtcgcgattatctatgatacacacagatgaaacactttaaagctcgt	1644
Db	1261	ggtatggaagtacaaaaaagtcgcgattatctatgatacacacagatgaaacactttaaagctcgt	1320
QY	1645	gaacactttagatgaatcgaagaactcttggagaccttactttagtgggaagtctggagac	1704
Db	1321	gaacactttagatgaatcgaagaactcttggagaccttactttagtgggaagtctggagac	1380

QY	1705	acacgtcgtgatataatttaagaatcaagcagcagcaatataacatctacccttaagga	1764
Db	1381	ACACGTGTGATTAAATTTTAAGAAATCAAGCAGACACATATTAACATCTACCCCTCACGGA	1440
QY	1765	atcactatgtccgtccctctgttatctcaaggaggtlaccaaaagcgtgttaaacattcgaag	1824
Db	1441	ATCACTGTATGTCCTCTCTCTTGTATTAAGAGATTTACCAAAGGTGTAAACATTTTGAAG	1500
QY	1825	gatttccaattccgcagcagagaataatccaatacaaatggacagtgtcgtcaagaat	1884
Db	1501	GATTTTCCAATTCTGCCAGGAGAAATTTTAAATATTAATTTGGCAGTGTACTGTACAAAGAT	1560
QY	1885	gggcccaactaaatcagaatccctcgtgtccgtacccgcattactctagtcttcgttaaatg	1944
Db	1561	GGGCCCAACTAAATCAGATTCCTCGGTGCTTACCCGCTTATTAACCTAGTTCTGTAAATAG	1620
QY	1945	gagagagatctagctttaaggaacctattggccctctctcatcttgcttacaagaagactgta	2004
Db	1621	GAGAGAGATCTTAGGCTTGAAGACTCATTTGGCCCTCTCTCATCTGCTACCAAAATCTGTA	1680
QY	2005	gattcaaaagaagaaacccagataatgtcagacaaagaagatgtacatccgtgtttctgtaltt	2064
Db	1681	GATCAAAAGAGAAACCGAATTAATGTACGAAACAAAGGAATGTATCTGTTTCTGTATT	1740
QY	2065	gatggaacccgaagctgtgtactccacagagaatatacaagcttctccccaatccagct	2124
Db	1741	GATGGAACCCGAACCTGTACTCCACAGAAATATACAAAGCTTTCTCCCCAATCCAGCT	1800
QY	2125	ggaggtgcagctcttgaggtaccagagttccaaagctccaaacatatgtcaacagatcaaatggc	2184
Db	1801	GGAGGTGCAGCTTGGAGATCCAGAGATTCCAAGCCCTCAACATATATGACAGCTCAATAGGC	1860
QY	2185	tatgtttttgtagttgtccagttgtcaagttgtgttcagagagttgcatacctgtgtacatt	2244
Db	1861	TATGTTTTGTAGTTTGTTCGACTGTCTCAGTTGTGTTTCATATAGAGTGGCATACGTGTACATT	1920
QY	2245	ctaagcatctggagacagacagactgacttcccttctgtcttctctctctgtgatataccctcaaa	2304
Db	1921	CTAAGCATTTGGAGACACAGACTGACTTCCCTTCTGTCTTCTCTGTGATATACCTTTCAA	1980
QY	2305	cacaaatagtctctatgaaagacaacatccactatcccatcttcaggagaaactgtcttc	2364
Db	1981	CACAAATAGTCTTATGAAAGACACACTCCCTTTTCCATTTCTCAGGAAACCTGTCTTC	2040
QY	2365	atgtcagatggaanaaccccaagttcatatgattctcgtgggtgcacaacatccaaacttcggagac	2424
Db	2041	ATGTGATGAGAAACCCAGGCTATATGATTCTTGGGTGTCCACACTCAGACTTTTGGAGAC	2100
QY	2425	agagagcatgaccgcgctactctgtaaggtttctagtctgtgacaagaagaactgtgtatattac	2484
Db	2101	AGAGGCATGACCCGCTTACTGAAGGTTTCTAGTTGTGACAAAGACACTGTGATTATTAC	2160
QY	2485	gaggaacggttatgaaagatatlttaagcatctgtcgtgtgtcaaaaaacaatgycattgtaacca	2544
Db	2161	GAGGCACTTTATGGAAGATATTTTAGCATATCTTGCTGATTAACAAACAAATGCCATTTTAAACA	2220
QY	2545	agaagctctcccaagaattcaagaacacccctagactagcacaagaagaaatlaatgtccacc	2604
Db	2221	AGA-----	2223
QY	2605	ccaccagctcttgaaacgcatacaacgggaataaactcgtactactcttcagtcagatcaaa	2664
Db	2224	-----GAAATTAACTCGTACTACTCTTCAGTCAAGATCAAA	2256
QY	2665	gaggaaatctgactatgtatgataccataatcaagtgtgaatgaaagaaagatttggacatt	2724
Db	2257	GAGGAATTTGACTATGTATGTATACCATATCATAGTTGGAATTAACAAAGGAATTTTGTACATT	2316
QY	2725	tatgatatagatgtaaaatcaagagcccccagcttcaaaagaanaaacacgcacatttt	2784
Db	2317	TATGTATGGATGTAAATATCAGAGCCCCGAGCTTTCAAAABAABAAACGACACTAATTTT	2376
QY	2785	attgcttcaggtggagagcctctgggaattatggtatgtagctgtcccaatgcttcttaaga	2844



|||||  
Db 2377 ATTGGTCAGTGGAGGCTCTGGGATTATGGATGATAGTACGCCCATGTTCTAAGA 2436  
Qy 2845 aacaaggctcaagatgagcagtgctccctcagtlcaagaagatgttttccaggaattacc 2904  
Db 2437 AACAGGGCTCAGAGTGGCAGTGTCCCTCAGTTCAAGAAAGTTGTTTCCAGGAATTTACT 2496  
Qy 2905 gatgctccttactcaagcctctataccgttggagaaactaataagacatttggagctctgt 2964  
Db 2497 GATGCTCCTTACTCAGCCCTTATACCTGGAGACTTAATGAACATTTGGGACTCTCG 2556  
Qy 2965 gggccatatataagagcagaagtltgaagataataatcattgaacttccagaatcaggcc 3024  
Db 2557 GGGCATATATAGAGCAGCAAGCTTGAAGATATCATCTGTAACCTTCAGAAATTCAGGCC 2616  
Qy 3025 tctcgtccctatctctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 3084  
Db 2617 TCTGTCCTCTATCTCTTCTATCTCTAGCCCTTATTTCTTATGAGAGACATCAGAGCAAG 2676  
Qy 3085 gcagaacctagaaaaaactltgtcaagcctaa tgaaccacaaacttacttltgaaagtgt 3144  
Db 2677 GCAGAACCTAGAAAAAATTGTTCAAGCCTAATGAACCAAAACCTTACTTTGGAAGTG 2736  
Qy 3145 caacatcatatgagcaccacataaagatgatttgaactgcaaaagcctgggctatattctct 3204  
Db 2737 CAACATCATATGAGCAGCCCACTAAAGATGATTGACTGCAAGGCCGCGCTTATTTCTCT 2796  
Qy 3205 gatgttgaccttgaaaaaagatgtgacacagccttgatctggagccctctgtgtctgacac 3264  
Db 2797 GATGTTGACCTTGAAAAAGATGTGCACCTCAGGCTGATTTGGACCCCTTGTGCTCCAC 2856  
Qy 3265 actaacacactgaacccctgtgtcaltgaggagacaagtgcagtlacagaaatttgcctgttt 3324  
Db 2857 ACTAACACACTGAACCCCTGCTCATGAGAGACAAGTGACAGTACAGAAATTTGCTGTGTT 2916  
Qy 3325 t---tcaccatcttltgtagagaccacaaagctgtgtacttcaactggaataatgtgaagaac 3381  
Db 2917 TTCCCTCACATCTTTGATGAGAGCAAAAGCTGTACTTCACTGAAATATGGAAGAAAC 2976  
Qy 3382 tgcagggctccctcgaatatcagaatgagaaatcccaattttaaagagattatgcttc 3441  
Db 2977 TGCAGGGCTCCCTGCAATATCCAGATGGAAGATCCCACTTTTAAAGAGATTTATGCTTC 3036  
Qy 3442 catgcaatcaatlygctacataa tgaatacactcctgtgtgttaatlygctcagaatcaa 3501  
Db 3037 CATCAACATCAATGGCTACATATATGATACACTGAGCTGTAGTATGCTCAGATCAA 3096  
Qy 3502 aggatltgattgtatctgtctcagcatlyggcagcaatgaaacatcattctatctatcttc 3561  
Db 3097 AGGATTCAGATGATCTCTCTCAGCATGGGCACCAATGAAACATCCATTCTATTCATTTTC 3156  
Qy 3562 agtgaacatgtgtcactcgtacgaaaaaaagagagatataaatgcaactgtacacatc 3621  
Db 3157 AGTGACATGTGTTCACGTACGAAAAAAGAGAGTATAAAATGGCACTGTACAAATCTC 3216  
Qy 3622 tatccaaagtlttlttgagacagltgaaatlytaccatccaaagctlygaaatttgcgcgtgt 3681  
Db 3217 TATCCAGGTGTTTGTGAAGACAGTGAATGTACCATCCAAAGTGAATTTGGCGGGTG 3276  
Qy 3682 gaatgacctatlttgagacatctacatgtgtgagatgagacaccttlttctgtgttaacag 3741  
Db 3277 GAATGCCCTTATTTGGAGCATCTTACATGCTGGATGAGCAGCACTTTTCTGTGTATCAC 3336  
Qy 3742 aataagttcgaactccctctggaaatgagcttcttgacacatagagatttcaagatttaca 3801  
Db 3337 AATAAGTGTCAAGACTCCCTTGGAATGGCTTCTGCACACATTAGAGATTTTTCAGATTACA 3396  
Qy 3802 gcttcagaacaatatgacagltggccccaaagcttgcacagacttcaatctacatccgatatca 3861  
Db 3397 GCTTCAGAGCAATATGACAGTGGGCCCCAAAGTGGCCAGACTTCATATTTCCGGATCA 3456  
Qy 3662 atcaatgctgtgagcacaagagagcccttctcttgatcaaggtgagctgttggacaca 3921  
|||||

Db 3457 ATCAATGCCGTGAGACCAAGAGACCCCTTTCTTGATATCAAGTGTATCTTTGGCACCA 3516  
Qy 3922 atgattattcaagcagcaatcaagaccaggggtgccgttcaagaatttccagcctctacatc 3981  
Db 3517 ATGATTATTACGGGCTATCAAGACCAAGGTTGCCCGTCAAGAAATTTTCCAGCTCTACATC 3576  
Qy 3982 tctcaagttatcatcatgatatagcttltga tggaaagatlygcagacttalcaggaat 4041  
Db 3577 TCTCACTTATCATCATGATGATAGTCTTGAAGAAAGTGCAGACATTATGAGAAAT 3636  
Qy 4042 tccactggaaccttaatgtgttcttlttgcaatlygagatcatctcgggataaacaacat 4101  
Db 3637 TCCACTGGAACCTTAATGAGTCTTTTGGCAATGTGATTCATCTGGGATTAACACAT 3686  
Qy 4102 atttttaaccctccaaatttltgctgatacatccgttgcacccaactatataagat 4161  
Db 3697 ATTTTAACTTCGCAATTTATGCTGATCATCCGTTTGCAACCACTTATATAGCAT 3736  
Qy 4162 cgaagcactcttcgaatlygagltgagtggtgtgaatlttaaataglttgacatgccaatg 4221  
Db 3757 CCAGACACTTTCGATGAGATGAGTGTGAGTGTGATTTAAATATGTCAGCATGCGCATTG 3816  
Qy 4222 ggaatlygaagtaagaacaatatcagatgcagatattcgtctcctacttaccat 4281  
Db 3817 GGAATGAGAGTAAGCAATATCATATGACAGATTACTGCTTCATCTCTACTTTAACCAAT 3876  
Qy 4282 atgttgcacactgtgtctcctcctcaaaagctgcactcactcccaaggaggaatgac 4341  
Db 3877 ATGTTGCCACCTGCTCTCTTCAAAAGCTGACTTTCACCTCCAAAGSAGAGATATGCC 3936  
Qy 4342 tggagacctcaglygaataatccaaagatlygctgcagaatlygacttccagaagaatg 4401  
Db 3937 TGGAGACCTCAGAGTAATATCAAAAGAGTGGCTGCAAGTGGACTTCCAAAGACAAATG 3996  
Qy 4402 aaagtccagggatcaactcactcagggagtaaatctcgtctaccagatatgtatgaag 4461  
Db 3997 AAAGTACAGGAGTAATCTCAGGAGTAATAAATCTGCTTACAGCATATATGTGAAG 4056  
Qy 4462 gagttccatctcagcagcagtaagaatggccatcagtgactctcttlttccaagaatgac 4521  
Db 4057 GAGTTCCTATCTCCAGCATCAAGATGGCCATGAGTGTGAGCTCTTTTTCAGATATGCC 4116  
Qy 4522 aaagtaaaagttlttccaaggaaatcaagaactccttcaacactgtlygtgaactcttagac 4581  
Db 4117 AAAGTAAAGTTTTCAGGGAATCAAGACTCTTCAACACGTTGTAACCTCTAGAC 4176  
Qy 4582 ccacgttactactcgtcacttctgaatttgaatcccccagagtttgggtgaccagaattgac 4641  
Db 4177 CCACCGTTACTAGCTGCTGCTGGAATTCACCCCAAGATGTGGTGCACCAATTTGCC 4236  
Qy 4642 ctgagatlygaagttctgtgctgcagagcaagagacctaac 4683  
Db 4237 CTGAGGATGAGGTTCTGTGGCTGCAGAGCAGACGACCTCTAC 4278  
RESULT 4  
108643 4281 bp PAT 02-DEC-1994  
LOCUS 108643  
DEFINITION Sequence 3 from Patent WO 8800831.  
ACCESSION 108643  
VERSION 108643.1 GI:588649  
KEYWORDS  
SOURCE  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 4281)  
AUTHORS Pasek, M.P.  
TITLE DNA SEQUENCES CODING FOR MODIFIED FACTOR VIII:C AND MODIFIED FACTOR VIII:C-LIKE POLYPEPTIDES AND PROCESSES FOR PRODUCING THESE POLYPEPTIDES IN HIGH YIELDS  
JOURNAL Patent: WO 8800831-A 3 11-FEB-1988;  
FEATURES  
source Location/Qualifiers  
1..4281



/organism="unknown"  
BASE COUNT 1245 a 943 c 946 g 1147 t  
ORIGIN  
Query Match 87.0% Score 4147.2; DB 10; Length 4281;  
Best Local Similarity 97.9%; Pred. No. 0;  
Matches 4269; Conservative 0; Mismatches 3; Indels 90; Gaps 3;  
QY 328 gccacccagaatactactccctggtgtgacgttggaactgtcatggagctatatagcgaagatgcat 387  
|||||  
Db 4 gccaccggaagatctactcctgggtgacgttggaactgtcatggagctatatagcgaagatgcat 63  
QY 388 ctccgttagctgctctgtgacgcgaagatttcctccttagagtgctcaaaatcttccattc 447  
|||||  
Db 64 CTGGGTAGGCTGCTGTGGAGAGCAAGATTCTCTAGAGGCCAAATCTTTTCCATTTC 123  
QY 448 aacacctcagtcgtgtacaaagaactcgttltgtagaatcacggttccacttccaac 507  
|||||  
Db 124 AACACCTCAGTCGTGTACAAAGAGACTCTGTTGTAGAAATTCAGATTCACGATTCACCTTTTCAC 183  
QY 508 atgcctaagccaagggccaccttgatgtgtctgtatgcttaccatccaggtcgaagtc 567  
|||||  
Db 184 ATCGCTAAGCCCAAGGCCACCTGGATGGTCTGCTAGGTCCTACCATCCAGGCTGAGGTT 243  
QY 568 tatgatacagtgatcatctacacttaagaacatggtctccatccctgfcagttcattcagtc 627  
|||||  
Db 244 TATGATTCAGTGGTCACTTAACTTAAGAACATGGCTTCCATCTGTCAAGTCTTTCATGCT 303  
QY 628 gttagtgatcctacttgaaagctctctgaggagcttgaaatagtagatcagaaccagtc 687  
|||||  
Db 304 GTTGCTGATCTCTACTGGAAGGCTTCTGAGGAGCTGAATATGATGATCAGACCAAGTCAA 363  
QY 688 agggagagaagaagatgtataagctctccctcgtgtggaagccaatatactgtctggcagtc 747  
|||||  
Db 364 AGGAGAGAAAGATGATTAAGTCTTCCCTGAGGAGCCATACATATGCTGGCAGAGTC 423  
QY 748 ctgaaagaagatggtccaatggtcctctgaccacactgtgcttaccctctatatcttct 807  
|||||  
Db 424 CTGAAAGAGATGGTCCCAATGGCTCTGACCCACTGTGCTTACTTACTATATCTTTCT 483  
QY 808 catgtgacctggttaaaagacttgaaattcagagctcatctggagccctactagtagtata 867  
|||||  
Db 484 CATGTGACCTGTGTAAGAAAGCTTGAATTCAGGCTCATTTGAGGCCCTACTAGTATGAGA 543  
QY 868 gaaggaggtctggtccaaagagaagacacagacttgcacaaatttatactacttctgtc 927  
|||||  
Db 544 GAAGGGAGTGTGGCCAAAGAGACACACACCTTGACAAATTTATATACCTTTTGTCT 603  
QY 928 gtaattatgaaggaaagatttgacactcagaaacaaagaactcctgtatgcagatag 987  
|||||  
Db 604 GATTTTATTAAGGAGAAATTTGGCACTCAGAAACAAAGAACTCTTGATGCGAGATAGG 663  
QY 988 gatgtgcatctgctcggcctggtcctaaatgcacacagtcacatgtgtatgttaaacag 1047  
|||||  
Db 664 GATGCTCATCTGCTGGGCTGGCCCTAAATGACACACACTCATATGTTATGTAACAGG 723  
QY 1048 tctctg----ccaagttcgtatgtatgtccacaggaatcactatgtgcatgtgattgga 1104  
|||||  
Db 724 TCTCTGCTACCAAGTGTGATTTGATGTCACAGAAATCACTTATTTGCGATGATTGGA 783  
QY 1105 atgggacacactcctgaagtgacactcaatatctcctgaaggtcacacattctgtgagg 1164  
|||||  
Db 784 ATGGGACACACTCTCTGAAGTGCATCAATATCTCTCGAAGGTACACATTTCTTTGAGG 843  
QY 1165 aacacatgcagagcgtctcttgaaatctgcacaataaacttctccttactgtcctcaaacatc 1224  
|||||  
Db 844 AACCATGGCAGGAGTCTCTGGAATCTCGCAATATCTTCCCTACTGCTCAACACCTC 903  
QY 1225 ttgatgaccttggagagcttctactgttttggcatatactcttcccaacaaatgatggc 1284  
|||||  
Db 904 TTGATGAGACCTTGAGACGTTTCTACTGTTTGTGATATCTCTTCCACCAACATGATGGC 963

QY 1285 atggaagctatgtcaagaagtagacagctgtccagagaagcccaactacgaatgaaat 1344  
|||||  
Db 964 ATGGAAGCTTATGTCAAGATAGACAGCTGTCCAGAGAACCCCACTACGATTAATAAAT 1023  
QY 1345 aatgaagaagcggagaagcatatgatgatcttaccgattcttgaaatgtatgtgtcag 1404  
|||||  
Db 1024 AATGAAGACGGGAAGACTATGATGATGATCTTACTGATTTCTGAATGATGTGTGCAGG 1083  
QY 1405 ttgatgatgaaactctctctctcttcttatacctaaatctgctcagttgtccaagaacatcc 1464  
|||||  
Db 1084 TTTGATATGACACCTCTCTCTCTTATCCAAATTTGCTCAGTTGSCAAGAAACATCTCT 1143  
QY 1465 aaaactggtgtacatactacatctgctcgtcgaagaagagcgtggaactatgctccctatgc 1524  
|||||  
Db 1144 AAAACTTGGGACTTATTCATTGTCTGTGAAGAGAGAGACTGGGACTTGTCTCTTATGTC 1203  
QY 1525 ctgcgcccgatgacagaagttataaagaatcaatattgaaacaaatggccctcagcgatt 1564  
|||||  
Db 1204 CTGCCCCCGATGCACAGAGTTATAAAGTCAATATTTGAAACATGSCCTCAGCGATT 1263  
QY 1585 gtttaggaagtacaaaaaagtcgcgaattatgtgcatacagaatgaacactttagatcgt 1644  
|||||  
Db 1264 GTTAGGAAGTACAAAAAGTCCGATTTATGGCATACACAGATGAACCTTTAAGACTCGT 1323  
QY 1645 gaagctatcagcatgatacagaagatcttggacacttacttcttggggaagttggaagc 1704  
|||||  
Db 1324 GAAGCTATTACAGCTATGATAGAGATCTTGGACCTTTACTTTATGGGAGATTGGAGAC 1383  
QY 1705 acaactgttgatlatatlttaagaatcaagcaagcagacataatacacttaccctcagga 1764  
|||||  
Db 1384 ACACGTTGATTTATATTAAAGATCAAGCAAGACGACATATACATCTACCCGACAGGA 1443  
QY 1765 atcaatgagtcgcgtctctgtatctcaaggaagattccaaaaggtgtlaaaacatttgaag 1824  
|||||  
Db 1444 ATCAGTATGTCCGTCTCTTGTATTAAGAGATTTACCAAAAGGTGTAAACATTTGAAG 1503  
QY 1825 gatttccaatctctgcccagagaagaatatacaaatataaaggaagtgatctgtagaagat 1884  
|||||  
Db 1504 GATTTTCCAAATCTGCCAGAGGAAATTTCAATATATAATGGACAGTGTATGAAAGAT 1563  
QY 1885 gggccaaactaaatcaagatctcgtgctgaccgctacttacttagttctgttaatatg 1944  
|||||  
Db 1564 GGGCCAACTAAATCAGATCTCGGTGCTGACCCGCTATTATCTGTAATGTTGTAATATG 1623  
QY 1945 gagagagatctagcttcaaggaactatggccctctcctactcgtctcaagaagatctgta 2004  
|||||  
Db 1624 GAGAGAGATCTAGCTTCAAGACTATGAGCCCTCTCTCATCTGTCTCAAAAGAAATCTGTA 1683  
QY 2005 gatcaagaagaagaacagataaagtcacagaagaagatgtcaatccgttttctgtaattc 2064  
|||||  
Db 1684 GATCAAAAGAGGAAACCGATATATGTCTGACAGAGAAATGTCAATCCCTGTTTCTGTATTT 1743  
QY 2065 gatgagaacccaagctgtgtacccctcaagagaatatacaacgcttctccccaatccaagt 2124  
|||||  
Db 1744 GATGAGAACCGAAGCTGTGCTCTACAGAGAAATATACAAAGCTTTTCCCAATCCAGCT 1803  
QY 2125 ggaagtgcagcttgaagatccaaggttccaagctccaactcaatgtcacagatcaatggc 2184  
|||||  
Db 1804 GGAAGTGCAGCTTGAAGATTCAGAGTTCCAAAGCCTTCCAAATCATATGCAACATCAAAAGGC 1863  
QY 2185 tatgttttgatagttgtgacgtgtcagttgttcttgataaggttggaactgtgatact 2244  
|||||  
Db 1864 TATGTTTGTATAGTTTGATGAGTTGTCAAGTTGTTTGTGATAGGTGCTATCTGTATATT 1923  
QY 2245 ctaagcatggaagcaagactgtacttcttctgtcttcttctcttgatataccttcaaa 2304  
|||||  
Db 1924 CTAAGCATTTGAGACACAGACTGACTTCTTCTGTCTCTTCTGTAATATCTTCAAA 1983  
QY 2305 cacaanaatgtctatagaagcacaatcaccctatcccatctcagaagaaactgtcttc 2364  
|||||  
Db 1984 CACAAAATGTTTATGAAGACACACATCCCTATTTCCCATTTCTCAGAGAAATGCTTTC 2043  
QY 2365 atgtcagtgtaaaacccaagttctatgtatcttctgtgggtgtccacaactcagacttctggaac 2424



|||||  
Db 204 ATGCGATGGAACCCAGGCTATGATTTGGGGTCCACAACTGACACTTTGGAAC 2103  
Oy 2425 agagcctgacgccttactaaagttctagctgagcaagaacacggtgattatc 2484  
|||||  
Db 2104 AGAGGCAATGACCGCTTACTAAGGTTTCTAGTTGTGACAAAGACACGTGATTAATTC 2163  
Oy 2485 gagagcactatgaagatattcagcacttctgagtaaaaaacatgccaatgaaca 2544  
|||||  
Db 2164 GAGGACAGTTATGAAGATTTTCAAGCATCTTGCTGAGTAAACAAATGCCATTGACCA 2223  
Oy 2545 agaagcttccccaattcaagacacctagcactagcgaagaatttaatgcaacc 2604  
|||||  
Db 2224 AGA----- 2226  
Oy 2605 ccacagctcttgaacgccatcaacgggaanaactgtaactcttcagtcagtcacaa 2664  
Db 2227 -----GAAATAACTCGTACTCTTCAGTCAGATCAA 2259  
Oy 2665 gaggaatctgactatgatacatalcatcagctgaagaagaagaatttgacatt 2724  
|||||  
Db 2260 GAGGAATTTGACTATGATGATACCATATCAGTTGAATGAAGAAGATTTTGACATT 2319  
Oy 2725 tatgataagatagaataatcaagcccccgacgttcaaaagaacacgacactattt 2784  
|||||  
Db 2320 TATGATAGGATGAATAATCAAGGCCCGCAGCTTTCAAAAGAAACAGACACTATTTT 2379  
Oy 2785 atgactcagtgagagagccttgatataatgagatgagtaagctcccaatgcttaaga 2844  
|||||  
Db 2380 ATTGCTCAGTGGAGAGAGCTTGAGATTATGAGATGATAGTACTCCCAATGTTCTAGA 2439  
Oy 2845 aacagagcctcagagtgagcagtgccctcactatgaagaagaattgcttccagaattact 2904  
|||||  
Db 2440 AAGAGGCTCAGAGTGGCAGTGTCCCTCAGTTCAGAAAGTGTTCAGGAATTTACT 2499  
Oy 2905 gatgagccttactacgacctatccgtggagaaactaaatgaaactttgagactcgt 2964  
|||||  
Db 2500 GATGGCTCTTACTACGCCCTTATACCGTGAGAACTAAATGAACATTGGGACTCCG 2559  
Oy 2965 ggagcatalaagaagcagaagttgaagataatcatcagtaacttcaagaatcaggcc 3024  
|||||  
Db 2560 GGGGCATATATGAAGACGAGAGTTGAAGATATATCATGTAATCTTCAGAAATGAGCC 2619  
Oy 3025 tctcgtccattcctctcattctagccttattcttaagaagaatcagaagcaga 3084  
Db 2620 TCTCGTCCCTATCTCTTATTTAGCCTTATTTCTTATGAGAGATCAGAGCAAGGA 2679  
Oy 3085 gacgaacctgaanaaacttgtaagcctaagaactaaacaaacttactttggaagt 3144  
|||||  
Db 2680 GCAGAACTTGAATAAACTTGTCAAGCTATATGAACCAAAACTTACTTTGGAAAGTG 2739  
Oy 3145 caacatcataatgacccaactaaagatgagcttgacgcaagccttgagctattcct 3204  
|||||  
Db 2740 CAACATATATGACACCCACATGAAGATGAGTTGACTGCACCAAGCCTGGCTTATTTCTCT 2799  
Oy 3205 gatgtgacctggaanaaagatgtaactcagcctgattgagccccctctgctgcac 3264  
|||||  
Db 2800 GATGTTGACCTGGAATAAGATGTGCACCTCAGGCTGATGGAACCCCTCTGCTGCCAC 2859  
Oy 3265 acataacacttgaacctgtctcagtgagagagaatgacagtaacgaagaattgctctgtt 3324  
|||||  
Db 2860 ACTAACACACTGAACCTGTGCTGATGGAGACAGTACAGTACAGGAATTTGCTGTGTT 2919  
Oy 3325 t-----tcaacatcttgaatgaaccaaagctgtaactcactgaanaatlatgaaagaac 3381  
|||||  
Db 2920 TTGCTCAGCATCTTTGATGAGACCAAAAGCTGTAATCTTACTGAAATATGGAAGAAAC 2979  
Oy 3382 tgcagaggtccctcgaatalatcagaatggaagatcccaactttaagaagaattatgccttc 3441  
|||||  
Db 2980 TGCAGGCTCCTCTCATATATCCAGATGGAAGATCCCACTTTTAAAGAAATATATGCTTC 3039  
Oy 3442 catgcaatcaatgctatcataatgataactactgctgaatgaatgctcagaatcaa 3501  
|||||

---

Db 3040 CATGCATCAATGCTACTATATGATATCACTACTGGCTTAGTAATGCTCAGATCAA 3099  
Oy 3502 aggattcgatglatcgtcgaacatgagcagaatgaanaacatcattatcatttc 3561  
|||||  
Db 3100 AGGATTGATGATCTCTCTCAGCATGGGAGCATGAAACATCAATGTGATTCATTTC 3159  
Oy 3562 agtgcacatglttcaactgtacgaaaaaagagagatataaatgacatgtaacatc 3621  
Db 3160 AGTGACATGTGTTCACTGTAAGAAAAAGAGATGTAATAAAGGACGTACAAATCTC 3219  
Oy 3622 tatccagtgcttttgaacacagtggaagtgttaccatcccaagcttggaatttggcgggtg 3681  
Db 3220 TATCCAGGTGTTTTGAGACAGTGAATGTTACATTCACAAAGCTGGAATTTGGCGGCTG 3279  
Oy 3682 gaatgcttattgagacatclacatgcttggaatgagacacttcttcgtgttaagc 3741  
Db 3280 GAATGCTTATTTGGGAGACATCTACATCTGGATGACACACTTTTCTGCTACAGC 3339  
Oy 3742 aataagtgtaactcccttggaatggtcttgagacacatlaagatattcagatata 3801  
Db 3340 AATAAGTGTGACACTCCCTGGGAATGGCTTGAGACACATTAGAGATTTTCAGATTACA 3399  
Oy 3802 gcttcagacacatagacatgagcagtgggcccaagccttgccagacttcatlccgagta 3861  
Db 3400 GCTTCAGACATATGAGACAGTGGGCCCAAGCTGGCCAGCTTCATTTATTCGGATCA 3459  
Oy 3862 atcaatgctcgtgagcacaagagcccttcttcttgatcaagtgatgctgttgaccaca 3921  
Db 3460 ATCAATGCTGAGAGACCAAGAGCCCTTTCTTGATCAAGTGGATGCTGTGGACCA 3519  
Oy 3922 atgattatcagcagcatcaagaacccaggtgcccgttcagaagaattctcagccttaac 3981  
Db 3520 ATGATTATTCAGGCGATCAAGACCCAGGCTCCGCTCAGAAATTCCTCAGCCTCTACATC 3579  
Oy 3982 tctcagttatcatatgattagcttctgaatggaagaagtggagacttctcagaagaat 4041  
Db 3580 TCTCAGTTATCATATATATATAGTCTTGATGGAAGAAAGTGGACATTCAGAGAAAT 3639  
Oy 4042 tccacttgaaccttaatgcttctcttggcaatgltgatalcatctggataaacaacat 4101  
Db 3640 TCCACTGGAACCTTATATGCTGCTTTGGCAATGGATTCATCTGGATTAACACAT 3699  
Oy 4102 attttaaacctccaattatgcttcogataacatccglttgacccaactatlaagatc 4161  
Db 3700 ATTTTAAACCTCCCAATATTTGCTGATACATCCGTTTGCACCAATCTTATTAAGCAT 3759  
Oy 4162 cgaagcactcttgcagatgagtgatgagctgtaatttaaatagltgcagatgcaatg 4221  
Db 3760 CGCACACTCTTCGATGAGAGTTGAGGCTGTGATTAATTAATTAATTCAGCATGCCATTG 3819  
Oy 4222 ggaatggaagtaaaagcaataatcagatgacacagatlaactgcttcatcacttaccat 4281  
Db 3820 GGAATGAGAGTAAGAAAGCAATATCAGATGCACAGATTAATCTGCTTATCTTACCAAT 3879  
Oy 4282 atgtttgcacctggtctccttcaaaagctgcacttcaactccaagggagagtaatgcc 4341  
Db 3880 ATGTTTGCACCTGCTCTCTTCAAAAGCTGCACTTCACCTCCAAAGGAGAGTAAGGCC 3939  
Oy 4342 tggagacctcagtggaataatccaaagaagtgtgtcgaagtggagcttccagaagaatg 4401  
Db 3940 TGGAGCTCTCAGGTGAATATATCCAAAGAGTGGCTGCAGTGGGACTTCAGAGACATG 3999  
Oy 4402 aaatgcacaggaatgaactactcagaaggaatgaatctctgcttaccagcatgltgaaag 4461  
Db 4000 AAAGTCACAGAGATTAATCAACGAGATTAATAATCTGCTTACACAGATGATGTAAG 4059  
Oy 4462 gagttcctcactccagcagtgcaagaatgagcatagtgagactcttcttccagatggc 4521  
Db 4060 GAGTTCCTCATCTCCAGAGTCAAGATGGCCATCAGTGGAGACTCTCTTTTTCAGAAAGGC 4119  
Oy 4522 aaagtgaaggttttcagggaaatcaagactccttcaacactggtggaactctcagac 4581  
Db 4120 AAAGTAAGGTTTTTCAAGGAATCAAGACTCTTCAACCTGTGGTGAACCTCTTAGAC 4179  
|||||



QY	4582	ccaccgttactgac	ctcgctactccttcgaattca	ccccccagattgggtgcac	caatgccc	4641
Db	4180	CCACGGTACTGAC	TGCGGTACCTTGAA	TTTCCACCCCGAGAGTTGGGTG	CACACAGATG	CC 4239
QY	4642	ctgagagatgaggttctctgtgctg	cgagagcacaagaccctac	4683		
Db	4240	CTGAGAGTAGAGG	TTCTGTGGCTGCAGAGC	CACAGGACCTTAC	4281	
RESULT	5					
LOCUS	108642					
DEFINITION	108642	4548 bp				
ACCESSION	108642	Sequence 2 from Patent WO 8800831.	PAT		02-DEC-1994	
VERSION	108642.1	GI:588648				
KEYWORDS						
SOURCE		Unknown.				
ORGANISM		Unknown.				
REFERENCE		Unclassified.				
AUTHORS		1 (bases 1 to 4548)				
TITLE		Pasek M.P.				
JOURNAL		DNA SEQUENCES CODING FOR MODIFIED FACTOR VIII:C AND MODIFIED FACTOR				
FEATURES		VIII:-LIKE POLYPEPTIDES AND PROCESSES FOR PRODUCING THESE				
		POLYPEPTIDES IN HIGH YIELDS				
		Patent: WO 8800831-A 2 11-FEB-1988;				
		Location/Qualifiers				
		1..4548				
		/organism="Unknown"				
BASE COUNT		1344 a 1004 c 1002 g	1197 t	1	others	
ORIGIN						

Query Match	85.7%	Score 4085	DB 10	Length 4548
Best Local Similarity	95.1%	Pred. No. 0		
Matches 4325	Conservative 0	Mismatches 31	Indels 192	Gaps 3
OY 328 gccaccagaatactactctggtgctgacgtgacacgtcatggtgactatgcaagtat				387
Db 1 gccacacaaagatctactccttgggtgcactggaactggtatggtggactatgccaagtat				60
OY 388 ctcggtgacgtgctcgtgtagacgcaagatcttcctcctaagtgccaaaatctcttcacatc				447
Db 61 CTCGGTAGAGTCGCTGTGGACGCAAGATTTCCTCCTACAGTGGCCAAATCTTTCCATTTC				120
OY 448 aacacctcaagtcggtgtacaaaaaagactctgtttgtagaatccaaggttcacctttcaac				507
Db 121 AACACCTCAGTCGGGTACAAAAAGACTCTGTTGTAGAAATTCACGGATCACACCTTTTCAC				180
OY 508 atcgcctaagccaagccacccttgatgggtctgctcaggttcctacacatccagcctgaagtt				567
Db 181 ATCGCTAAGCCAAAGGCCACCTCGATGGGTCTGCTAGTCTCTACCATTCACAGCTGAGGTT				240
OY 568 tatgatacagtggtcatctatacctaaagacaatggcttcccatcctgtlcaagcttcatagt				627
Db 241 TATGATACAGTGGTCATTTACACTTAAGAAACATGGGTTCCATCCGTCAGTCTCATGCT				300
OY 628 gttgtgtatcctactactgtaaaagctcttgtagagagctgataatgatgatccagaacgtcaa				687
Db 301 GTTGGTGTATCTCCTACTGGAAAGCTTCCTAGGGAGACTGTAATGTAGATCAACACCACTCAA				360
OY 688 agggagaagaagaagtatgaatgaagctctccctcgttggaagcaatacatatgcttgcaagtc				747
Db 351 AGGGAGAAAGAGATGATTAAGTCTTCCCTGGTGGAAAGCCATPACATATATGTGGCAGGTC				420
OY 748 ctgaaagaagaatggtlcccaatggcctctgacccactggtgcttactactactaatacttct				807
Db 421 CTGAAGAGAGATGGTGCATGAGCCCTGACCCACAGTGTCTTACTACTCATATCTTCT				480
OY 808 catgtgagccctggtataaagaacttgaaatccaggtctcatatgtagccctactagtatgtaga				867
Db 481 CATGTGAGCCCTGTATAAAGACTTGATTCACGGCCATCTGAGGCCCATCTGATGTGTGTGA				540

QY	866	gaaggaagatctgcggccaaggaagaacacagactctgcacaatttatactacttctgtc	927
Db	541	GAAGGAAGTCTGGGCCAAGGAAAGACACACACCTTGCCAAATTTATCTACTTTTGGCT	600
QY	928	gtaatttgtagaagggaaaagtctgacatcaagaacaaagaaccccttggtagaagatag	987
Db	601	GTAATTGTAGAAAGGAAAGTAAGTTGGCATCTCGAAGAACAAAGACCTCTTGATGACGAGTATGG	660
QY	988	gatctgatcatctgcggccctgcggccctaaagtacacacagatcaattgltatglaacag	1047
Db	661	GATCGTGCATCTGCTCGGGCCCTGGCCCTAAATATGCACACAGTCATATGTTATGTAAACAG	720
QY	1048	tctctg---ccaagtcctgatgtgatgcccacaggaatacagtcatactgacatglatgga	1104
Db	721	TCTCTGTACCAACAGCTCGATGTGATGGCCACAGGAAATACGTCTATTTGGCATGTGATGTGA	780
QY	1105	atgggcaccactcctctgaagtgacatcaattctcctgaaggtacacacatttcttggag	1166
Db	781	ATGGGCACACCTCTCGAAGTCACATCAATTCCTCGAAGATCACACATTTCTTGAGG	840
QY	1165	aaccaatgcacagcgctcccttggaatctgcacaataacttccctctgcctcaaacac	1222
Db	841	AACCATGCCCAAGGGGCTCCTTGGAATATGCGCCATTAATCTTCTTACTCTCAACACCTC	900
QY	1225	ttgatvgaccttggacagtttctactcgttttgatcatatcctctcccaacaatgatgac	1284
Db	901	TTGATVGACCTTTGACAGTTTCTACTGTTTGTCATATCTCTTCCACCAACATGATGGC	960
QY	1285	atggaagcttatgtgcagaatgagaagctgtgccagggaaaccccaactacgaatgaaaat	1344
Db	961	ATGGAAGCTTATGTGCAAAAGTAGACAGCTGCCAGAGAACCCCACTACGAATGAAAT	1020
QY	1345	aatggaagaagggagaagaciatgatgcagatccttaccgatcttgaaatggatgtgtcag	1404
Db	1021	AATGAGAAGCGGGAAGACTATGATGATGATCTTACTGATTTGAAATGGAATGGATGGCTAG	1088
QY	1405	tttgaatgaacacatcctcccttctatccaatctgcgtcagttgcgaagaagacatc	1466
Db	1081	TTTGATGATGACAACTCTCCTCTTATCCAAATTTGCTCAAGTTCGCAAGACATGCT	1140
QY	1465	aaaacttgggttaacttatcatctgcctgcctgaaagaagagacttggagactatgctccctatgc	1522
Db	1141	AAAAATTTGGTACATTTACATTGCTGCGAAGAGAGAGACTGGGACATATCTCCCTTAGTC	1200
QY	1525	ctgcgccccggaatgacaagaatgataaagaatcaatttgaacaatbgccctcaagggat	1584
Db	1201	CTCGCCCCGATGACAGAAAGTTAAAAAGTCATATTGAGACATATGGCCCTCAGCGGAT	1266
QY	1585	ggtagaagaatcaaaaaagtcgcgaatttatgycatacacagatgaaacctttaaactcgt	1644
Db	1261	GGTAGAAGTACAAAAAAGTCCGATTATATGGCATACAGATGAACCTTTTAAGACTCGT	1320
QY	1645	gaagctatcagcatgaaatcaaggaaactctgggaaccttacttataggggaaatgtggagac	1704
Db	1321	GAAGCTATTTCAGCATGTGATAGGAATCTTGGGACCTTTACTTATGGGGAAGTTGGAGAC	1380
QY	1705	acaagttgatcatattttaagaatccaagaagaagacatataaactatcacctcacaaga	1764
Db	1381	ACACGTGTTGATTATATTTAAGAAATCAAGCAGACGACCATATATACATCTTACCTCTACGGA	1440
QY	1765	atcaactgatgtccgttcttgtatctcaaggagatltaccacaaggltglaaacaatttga	1824
Db	1441	ATCACTGATGTGCGCTCTTGTATTAAGGAGATTTACCAAAAGGTGTAAACATTTGAAG	1500
QY	1825	gattttccaattctgcgaaggaagaatatltcaaatataaatgagacgtgtacgtgtagaagat	1888
Db	1501	GATTTTCCAAATTTGCGCAGAGAAATATTTCAAAATATAATGTAGACAGTGCATCTGTAAAGAT	1566
QY	1885	gggccaactaaatcaagctccctgcgtgcctgaaccggctatacttaccctagtttggtaatg	1944
Db	1561	GGGCGCACTAATATAGATCTCTCGGTGCTGACCCGCTATTACTCTAGTTTCTGTTAATATG	1620
QY	1945	gagagaagctctagcttcaagacatcaattggccctctccatctatgctatcaagaagaactgta	2004







QY 3976 tacatctcagttatcatcatgtatagtccttgatggagaagatggcagactatcga 4035  
|||||  
Db 3841 TACATCTCTCATGTTATCATCATGTATAGTCTTGATGGAGAAAGTGGCAGACTTATCGA 3900  
QY 4036 ggaatctcactggaaacttaatggtctctcttgccaatgltgatalcatctggataaaa 4095  
|||||  
Db 3901 GGAATTTTCACCTGGAACTTAATGCTCTCTTGGCAATGGATTCATCTGGGATGAAA 3960  
QY 4096 cacaatattttaacctccaatattgtctcgatatacaccgttttgacccaactaatat 4155  
|||||  
Db 3961 CACAAATTTTTTAACTCCCAATTTATGCTGATACATCCGTTTGCCCAACCACTTAT 4020  
QY 4156 agcatctgcagcaactctcgatggatgtatggctgtgatttaaatgttgagaatg 4215  
|||||  
Db 4021 AGCATTTGCGACGACTCTTCCCATGGAGATTGATGGGCTGTATTTAAATAGTTGACGATG 4080  
QY 4216 ccatltyggaatggagatlaaagaacatatacagatgacagataactgcttcatctat 4275  
|||||  
Db 4081 CCATTTGGGAATGGAGATTAAGCAATATCAGATGCACAGATTACTGCTCATCTACTTT 4140  
QY 4276 accaatagtttgcccactggtctctcttcaaaagctcgacttcaactccaagggagagt 4335  
|||||  
Db 4141 ACCAATATGTTTGCACCTGGTCTCTTCAAAAGCTCGACTTCACCTCAAGGAGAGAT 4200  
QY 4336 aatggcctggagactcagtgatgaataatccaaaagatggtctgaagtgagactccaaga 4395  
|||||  
Db 4201 AATTCCTGGAGACCTCAGGTAATATCCAAAAGAGTGGCTGCAAGTGAGCTTCACAGAG 4260  
QY 4396 acaatgaatatacagagatgaactatacagagatlaaactctctgcttaccagatgat 4455  
|||||  
Db 4261 ACATGAATATCACAGAGATTAAGTACTACAGAGTAAATCTGCTTACACATGTAT 4320  
QY 4456 gttaaggaagttctcatalctccagcagatcaaatggtccatagtgagactcttcttctcag 4515  
|||||  
Db 4321 GTCAAGAGATTCCTCATCTCCACAGCAAGATGAGGCTCATAGTGAGCTCTTTTTCAG 4380  
QY 4516 aatggcgaatlaaagtttctcagggaaatcaagactccttcaactgtyggtgaactct 4575  
|||||  
Db 4381 AATGGCAATTAAGTTTTTCAGGGAAATCAAGACTCTTTCACACCTGTGGTGAAGTCT 4440  
QY 4576 ctgaaccacagcttactgactcgtactcctcgaatcaaccacagagtttggtygcacag 4635  
|||||  
Db 4441 CTGACCCACCGTTACTGACTCTGCTTCAATTCACCCCAAGATTGGTGGACACAG 4500  
QY 4636 attgacctgagatgagatgttctggtctcgagagcagaggaactcttac 4683  
|||||  
Db 4501 ATTGCCCTGAGATGAGGTTCTGGGCTGCGAGGACAGAGACTCTTAC 4548

RESULT 6  
LOCUS 108641 4551 bp PAT 02-DEC-1994  
DEFINITION Sequence 1 from Patent WO 8800831.  
ACCESSION 108641  
VERSION 108641.1 GI:588647  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 4551)  
AUTHORS Pasetk,M.P.  
TITLE DNA SEQUENCES CODING FOR MODIFIED FACTOR VIII:C AND MODIFIED FACTOR VIII:C-LIKE POLYPEPTIDES AND PROCESSES FOR PRODUCING THESE  
POLYPEPTIDES IN HIGH YIELDS  
Patent: WO 8800831-A 1 11-FEB-1988;  
Location/Qualifiers  
1..4551  
source  
BASE COUNT 1345 a 1003 c 1004 g 1198 t 1 others  
ORIGIN

Query Match 85.7%; Score 4085; DB 10; Length 4551;  
Best Local Similarity 95.1%; Pred. No. 0;  
Matches 4325; Conservative 0; Mismatches 31; Indels 192; Gaps 3;  
QY 328 gccacagaagataactactcctggatgacagtggaactgtcatgaggaactatgcaaatgcat 387  
|||||  
Db 4 GCCACAGAAAGATATCTACTGCTGGGTGAGTGAAGTGCATGCGGAGCTATATGCAAGAT 63  
QY 388 ctggatgagctgctctggtggcgcgaagaattccctccctgaagtgccaaatcttctcattc 447  
|||||  
Db 64 CTCGGTGAAGTCTGCTGGACGCAAGATTTCTCTGAGCTGCAAAATCTTTTCCATTC 123  
QY 448 aacaccccaatgctgtatacaaaagactctgtttgtaaatctcaaggttcccttcaac 507  
|||||  
Db 124 AACACCTCATCTGCTGTAACAAAAGACTCTGTTGTAGAAATTCAGGATCACCTTTTAC 183  
QY 508 atcgtaagccaaagggccacccttgatggtctgtaagtgcttaccatccacagctgaagt 567  
|||||  
Db 184 ATCGCTAAGCCAAAGGCCACCCCTGGATGGTCTGTAGTCTTACATCCAGGCTGAGGTT 243  
QY 568 tatgatacagtggtcatataacttaagaacatggtctccatctcgtcagttctcatgct 627  
|||||  
Db 244 TATGATACAGTGTGATTAACCTTAAGAACATGGCTTCCATCCGTGCAGCTTTCATGCT 303  
QY 628 gttggtatctctacttggaaagcttctgaggaggtgaatatgataatgtaagaccagca 687  
|||||  
Db 304 GTTGGTATCTCTACTGGAAGGCTTCTGAGGAGCTGAATATGATGATGCAGACCAATCAA 363  
QY 688 agggagaagaatgataaagatcttccctggttggaaagccatacatgctgcaaggtc 747  
|||||  
Db 364 AGGAGAAAGATGATGATTAAGTCTTCCCTGTTGGAGAACCATATATGCTGGCAGGTC 423  
QY 424 CTGAAGAAGATGTGCTCAAGTGGCTCTGACCCACTGTCCTTACTATATCTTCTTCT 483  
Db 748 ctgaagaagaatggtccaatggtccttgaccacatggtcttactcatatcttctc 807  
QY 484 catgtgacctgttlaaagacttgaattcagggcttcatltagacccctacagatgata 867  
|||||  
Db 808 catgtgacctgttlaaagacttgaattcagggcttcatltagacccctacagatgata 867  
|||||  
Db 484 CATGTGACCTGTTAAAGACTTGAATTCAGGCTCATTTGGAGCCCTACTAGTATGTA 543  
QY 868 gaaggagatctggccaaagaaagacagacacttgcacaaattataactacttcttgc 927  
|||||  
Db 544 GAAGGAGTGTGGCCAGGAAGAACACACACCTTGCAAAATTTATATCTTTTGTCT 603  
QY 928 gttattgataaaggaagttggtcactcagaacaaagaacctcttgatgagcagatg 987  
|||||  
Db 604 GTATTTGATGAAGGGAAGATTTGGCACTGCAAAACAAAGATCTTGATGCAAGATAGG 663  
QY 988 gatgtgcatctgctcggtcgtggccttaaatgacacagtcaatggttataaagcag 1047  
|||||  
Db 664 GATGCTGATCTGCTCGGGCTTGCCCTTAAATGACACACATGATGTTATGTAACAGG 723  
QY 1048 tctctg----ccaggtctgattgattgacacaggaatcaatcagttctatlgcatgtgata 1104  
|||||  
Db 724 TCTCTGTACACAGGTCTGATTTGATGTCACAGGAATATCACTATATTGGATGATGGA 783  
QY 1105 atgggacacactcctgaatgacatcaatatctcctgaaggtacacacatttcttgag 1164  
|||||  
Db 784 ATGGGACCACTCTCTGAAGTGCACCTCAATATTCCTGCAAGGTCAACATTTCTGTGAGG 843  
QY 1165 aaccatgcagcagctccttggaatctgcacataacttcttactgtctcaaacctc 1224  
|||||  
Db 844 AACCATGCGCAGGCGTCTTGGAATCTGCCAATATCTTCTTACTGCTCAAAACCTC 903  
QY 1225 ttgatggacttggacagtttctactggtttgtcatatctcttccacacacatgagtc 1284  
|||||  
Db 904 TTGATGACCTTGACAGATTTCTACTGTTTGTGATATCTCTTCCACCAACATGATGCG 963  
QY 1285 atggaagcttatgtcaaaagtagacagctgtccagaggaagacccaactacagaaat 1344  
|||||  
Db 964 ATGGAAGCTTATGTCAAAAGTAGACAGCTGTCAGAGGAACCCCACTTACCAATGAATAAT 1023  
QY 1345 aatgaagaagcgaagaactatgatatgattcttactgattctgaatgagtggtcag 1404



```
|||||
Db 1024 AATGAAGAGGGAAGACATGATGATGATCTACTGATTTGGAATGATGATGCTCAG 1083
Oy 1405 ttgatgatgaacactcccttccttcttaccaaatcgtcagttgccaagaacatcct 1464
Db 1084 TTTATGATGCAACTCTCCTTCTTATCCAAATTCGCTGCTGACAGAGATCCT 1143
Oy 1465 aaacttggatcataltaacttgccttgctgaagagagaactgggaactaigtccctta 1524
Db 1144 AAAACTGGGATACATTATGCTGCTGAGAGAGACTGGGACTATGCTCCCTTAGTC 1203
Oy 1525 ctgcccccgatgacagaagtataaagaataattgaacaatggccctcagcgatt 1584
Db 1204 CTCCTCCCCGATGACAGAGATTAAAGTCATTTGAAACAAAGGCCCTCAGCGGATT 1263
Oy 1585 gttggaagatcaaaaaagtcgatttatggcatcacagatgaaacctttaagctcgt 1644
Db 1264 GGTAGAGATGCAAAAAGTCCGATTTATGGCATACAGATGAACCTTTTAGCTGCT 1323
Oy 1645 gaagctatlcagcatgaaatcaggaactctgggaaccttacttactatgggaagttagagac 1704
Db 1324 GAACGCTATTCAGCATGATGATCGAATCTTGGGACCTTACTTATGGGGAAGTTGAGAC 1383
Oy 1705 acaatgttgattataattagaatcaagaagaagccaatataactacccctcagga 1764
Db 1384 ACACGTGTGATTATATTTAAAGATCAACCAAGCACCATATMACATCTACCCTCACGGA 1443
Oy 1765 atcaatgatgcgctccttgatcacaagagattaccaaaagtggtaaacatttgag 1824
Db 1444 ATCCTGATGTCCTCCTCTTTGATTTCAAGGAGATTACCAAAAAGTGTTAAACATTTGAG 1503
Oy 1825 gatttccaactctgcagagaaataltccaataataatgagacagtgactgtagaagat 1884
Db 1504 GATTTTCCAATTTCTGCCAGAGAAATATTCAAAATTAATGAGACGTGACTGTAGAAAGAT 1563
Oy 1885 gggccaactaaatcaagatcccggtgcctgcgacctatactactcagtttgttaatatg 1944
Db 1564 GGGCAACTAAATCAGATCCCTGGGCTGACCCGCTATCTAGTTTCCTTAATATG 1623
Oy 1945 gagaagatcagcttcaagatcacttgccctccctccatctgtctaaagaactctta 2004
Db 1624 GAGAGATCTAGCTTCAGGACTCATTTGGCCCTCTCTCATCTGCTACAAAGAAATCTGTA 1683
Oy 2005 gatcaaaagagaacacagataatgtcagaacaaggaatgtcaatcctgttcttctatct 2064
Db 1684 GATCAAGAGAGAAACCGATATATGACAGAAAGAGATGCAATCTGTTTCTGTATTT 1743
Oy 2065 gatgagaacggaagctggttaacctcaagagaatatcaagcttctccccaatccagct 2124
Db 1744 GATGAGAACCGAAGCTGTACTCACAAGAAATATCAACGCTTTCCTCCCAATCCAGCT 1803
Oy 2125 ggaatgagcttggagatccagagtttccaaagctccaaacatcatgacagcatcaatggc 2184
Db 1804 GGAATGAGCTTGAGATCCAGAGTTCCAAAGCTTCCAAATCATATGCAACGATCATATGGC 1863
Oy 2185 tatgttttgataagtctgcagttgtlcaagttgttctgcataggagtgcaactaagtacat 2244
Db 1864 TATGTTTGTATAGTTTGTCAATGTGTACGTTGTTTGTGATGAGTGGCATACGTGATATT 1923
Oy 2245 ctaagacttggagacagactgacttctctgtcttctctctctcgtgataacttcaaa 2304
Db 1924 CTAGAGCTTGGAGCAGACAGACTTCCCTTCTGTCTTCTCTGAGTATACCTTTGAAA 1983
Oy 2305 caaaaaatgtctcatgagacacactcaactcaacttccactctcagaggaactgtcttc 2364
Db 1984 CAAAAATGCTTATGAGACACACTCACCTATTTCCATTTCTCAGGGAATCTTCTTC 2043
Oy 2365 atgtcgaatgaaaaacccagtgctcatggaatctgggtgtgcacaaactcagacttctggaac 2424
Db 2044 ATGTCATGGAACCCAGGCTATGATCTCTGGGGTCCCAACTCAGACTTTTGGAGAC 2103
Oy 2425 agaggaatgacgccttaactgaaagttctcagttgtgacaagaacactggtgatatctac 2484
|||||

|||||
Db 2104 AGAGCATGACCGCTTACTGAGAGGTTTCTAGTTGTGACAAAGAACTGCTGATTTATAC 2163
Oy 2485 gagaacagtatgaagatatttcaagcactctgtcgaataaacaatgcatatgaacca 2544
Db 2164 GAGGACAGTTATGAGATATTTTACGATACACTTTCGTGATTAACAAATGCCATTGAACCA 2223
Oy 2545 agaagctctcccaaat----- 2562
Db 2224 AGAAGCTTCCACAGATCTCTTCTGATTAACCACTATGTGACTCAGATACAAA 2283
Oy 2563 ----- 2562
Db 2284 GAAGGTGGAATTCOAAGAGAAGTCAACGAAAAACAGCTTTTAAGAAAAAGATACC 2343
Oy 2563 ----- 2562
Db 2344 ATTTGTCCCTGACGCTTGTGAAGCAATATGCAATAGCAGCAATTAATGAGGACAA 2403
Oy 2563 -----tcaagacacctagcactaggcaaaagcaattaat 2598
Db 2404 AATAAGCCCGAAATGAGATGACCTGGGCAAGCAAGTAGACTGAAGGCTGTGCTCT 2463
Oy 2599 gcaaccccaacagctcttgaacgcgcataacggaataactgttactactcttcagta 2658
Db 2464 CAAAACCCACAGTCTTGAAAGCCCATCAACGGAAATTAACCTCFACTCTTCACTCA 2523
Oy 2659 gatcaagagaatgtactatgatataccaatcagttgaatgaaagaagaagattt 2718
Db 2524 GATCAAGAGGAATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2583
Oy 2719 gacatttatgtagatgaaatcaagaagcccgcaagcttccaagaagaacacagcac 2778
Db 2584 GACATTTATGATGAGAGATGAAGAAATCAGAGCCCGAGCTTTCAAAAGAAACACGACAC 2643
Oy 2779 tatattatgtcgaagtgtgaagagctcgtggatgaigtatgagatgaactcccaatggt 2838
Db 2644 TATTTTATGTGCGAGTGAAGAGGCTCTGGGATTTAGGATGATGATGATGATGATGATGAT 2703
Oy 2839 ctaagaacaagagctcgaagtgagtgctccctcaagttccaagaagtgttctccaagaa 2898
Db 2704 CTAAAGAACAGGGCTCAGAGGCGAGTGCTCCCTCAGTTCAAGAAAGTGTGTTCCAGAA 2763
Oy 2899 ttactatggtctcccttactcagcccttataccgttggagaactaaatgaacatttggga 2958
Db 2764 TTTACTGATGCTCTCTTACTACAGCCCTTATACCGTGAGAACTAAATGAACATTTGGGA 2823
Oy 2959 ctccgtggccatatataagagcagaagtgtgaataataatcatatgtaactttcaagaat 3018
Db 2824 CTCTGGGGCCATATATATAGAGCAGAAAGTTGAAGATATATCATATGATCTTTCAGAAAT 2883
Oy 3019 caggcctctgctccctatctctctatctcagcttatcttcttataggaagaatcaaga 3078
Db 2884 CAGGCTCTGCTCCCTATCTCTTATTTCTAGCTTATTTCTTATGAGAAAGATCAAGG 2943
Oy 3079 caaggaagcagaacctagaaaaaaacttgtcgaagcctaaatgaacaaacttactttgg 3138
Db 2944 CAAGGACGACGAACCTAGAAAAAACTTTGTCAAGCCATATGAACCAAAACTTACTTTGG 3003
Oy 3139 aaagtgcacacatcatatagagcccaactaaagatgaagttttagatgcgaagcttgcctat 3198
Db 3004 AAAGTGCAACATCATATAGGACCCACCTAAAGATAGTTTGATGTCAAAAGCCTTG6GCTTAT 3063
Oy 3199 ttctctgaatgttgaacbtgaaaaaagatgtgcacccaagcctgatttgaacccctctgtc 3258
Db 3064 TTCTCTGATGTTGACCTTGAAAAAAGATGTGACACTCAGGCTGATTTGGAACCCCTTCTGTGTC 3123
Oy 3259 tggcaactaaacacactgaacctgtctcatatggagacaagtgtgacgtacagaaatgtgtc 3318
Db 3124 TGCCACACTTAACAACATGAAACCTGCTCATGTGAGACAAAGTACAGATACAGAAATTTGCT 3183
Oy 3319 ctgttttt---tcacacattttgatgagaccaaaagctggttaacttcaatgaaatatggaa 3375
Db 3184 CTGTTTTCCTACATCTTTGATGAGACCAAAAGCTGTGATCTTCACTGAAATAATGGA 3243
```



QY 3376 agaaatgcagagctccctgcgcaatataccagaatggaagatccaccattttaagaagaattat 3435  
|||||  
Db 3244 AGAAACTGCAAGGCTCCCTGCATATATCCAGATGGAAGATCCACTTTTAAAGAAATATAT 3303  
QY 3436 cgcctccatgcatacaatgagctacataatgatatacactactgagcttagtaatgagccag 3495  
|||||  
Db 3304 CGCTTCATGGAATCAATCAATGCTATATGATATGACTACTGCTGCTAGTATGATGATGAT 3353  
QY 3496 gatcaaaagatcgaatgatatctgctcaagaatgagcagaatgaaacatccattctatt 3555  
|||||  
Db 3364 GATCAAAAGATTCATGATGATGCTGCTGAGCATGGGCAATGAAACATCACTTCTATT 3423  
QY 3556 cattcagtggaactggttcacactgctcgaaaaaaagaagatataatggagctgtac 3615  
|||||  
Db 3424 CATTTAGTGAGCAATGATGCTTCTGACGAAATAAGAGATGATGAAATGAGCACTGTAC 3483  
QY 3616 aatcctacacaggtgcttcttgagacagtggaatggttacatccaaagctggaattcg 3675  
|||||  
Db 3484 AATCTCATCCAGGTGTTTTGAGACAGTGAATGTTACATCCAAAGCTGGAATTTGG 3543  
QY 3676 cgggtggaatgctcttattgagcagatcacaatgctgagatgagacacttcttggtg 3735  
|||||  
Db 3544 CGGCTGGAATGCTTATTGCGGACATCTACATGCTGGGATGAGACACACTTTTCTGCTG 3603  
QY 3736 taacagcaataagtgctgagactcccttggaatgagctcttgagacattagagatttcag 3795  
|||||  
Db 3604 TACAGCATATAGTGTCTGAGCTCCCTGGAAATGGCTTCTGGACACATTAGGATTTTCAG 3663  
QY 3796 attacagcttcagagacataatgagacagtgagcccaagaagctgagcagaactcattatcc 3855  
|||||  
Db 3664 ATTACAGCTTCAGGACAAATATGAGCATGGGCCCCAAAGCTGGCAACACTTCAATTATCC 3723  
QY 3856 ggaatcaatcagctcgtgagacacaaagagcccttcttgagatcaagtgagatcttg 3915  
|||||  
Db 3724 GGATCAATCAATGCTGAGACACCAAGAGCCCTTTCTTGATCAAGTGATCTGTG 3783  
QY 3916 gacacatgatttattacagcagcatcaagaagctgagcagagcttcagagcttcagagctc 3975  
|||||  
Db 3784 GCACCAATGATATTCAGGAGATTAAGACCAGGGTCCCTGCAAGATTTCTCCAGCTCC 3843  
QY 3976 taactcctcagttatcatcatcatgtaagtccttgatgagaaagatgagcagactatcga 4035  
|||||  
Db 3844 TACATCTCTCAGTTTATCATCATGATGATGATGATGATGATGATGATGATGATGATGATG 3903  
QY 4036 ggaatcagcagctgagacctaagtgctctctcttgagcaatgagatcattctgagataaa 4095  
|||||  
Db 3904 GGAATTCACACTGGACACTTAATGCTCTTCTTGCAATGATGATGATGATGATGATGATG 3963  
QY 4096 cacaatattttaacctccctcaaatattgctcagatatacagtttgagcacaacacattat 4155  
|||||  
Db 3964 CACATATATTTTAACTCCCTCCAAATATGCTGATACATCCGTTTGACCCCAACTCATATAT 4023  
QY 4156 agcatcgcagcagctctcgcagatgagatgagtgagcgtgataatlaaagtgtgagatg 4215  
|||||  
Db 4024 AGCATTTGCGACGACTCTTCGATGGAGTTATGATGGCTGTGATTAATATGATGATGATG 4083  
QY 4216 ccattggaatgagatgaaagaatatacagatgacagatctactgctcactactt 4275  
|||||  
Db 4084 CCAATGGGAATGAGATTAAGCAATATGATGATGATGATGATGATGATGATGATGATGATG 4143  
QY 4276 accaatatggttgcagctgctcctcctcaaaagctgagacttcaacagaggaagat 4335  
|||||  
Db 4144 ACCATATATGTTTGCACCTGGCTCTCTTCAAAAGCTGACTTCCCTCCAAAGGAGAGAT 4203  
QY 4336 aatgcctgagagactcagtggaataatccaaaagatgagctgagcagtgagctccaaag 4395  
|||||  
Db 4204 AATGCTTGAGAGACTTCAGGATTAATCAAAAGATGGCTGCAAGTGGACTTCCAAAG 4263  
QY 4396 acaatgaaatcacagagtaactactcagggagtaaatctctgcttaccagcatgtat 4455  
|||||  
Db 4264 ACAATGAAGATCACAGAGATTAATCAAGGAGTAAATCTGCTTACACAGCATGTAT 4323

QY 4456 gtgaagagatttcctcatctccagcagtcgaatgagcatcagatgagactctcttttcag 4515  
|||||  
Db 4324 GTGAAGAGATTCTTCATCTCCAGCAGTCAAGATGAGATGAGATGAGATGAGATGAGATG 4383  
QY 4516 aatgcaaaatgaaagatttctcaaggaaatcaaaagactcctcacacctgtgtgtaactc 4575  
|||||  
Db 4384 AATGCAAAAGTAAAGGTTTTTCAAGGAAATCAAGATCTCTCACACCTGATGATGATGATG 4443  
QY 4576 ctgaaccacacgttactgactgcttactccttcgaatcaaccccaagattggtgacag 4635  
|||||  
Db 4444 CTAGACCCACCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 4503  
QY 4636 attgccttgagagatgagatgctgtgctgagcagcaagagaccctac 4683  
|||||  
Db 4504 ATTGCTTGAGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4551  
  
RESULT 7  
AR034084  
LOCUS AR034084 5035 bp DNA PAT 29-SEP-1999  
DEFINITION Sequence 1 from patent US 5869292.  
ACCESSION AR034084  
VERSION AR034084.1 GI:5949689  
KEYWORDS  
SOURCE  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 5035)  
AUTHORS Voorberg,J.J.  
TITLE Hybrid proteins with modified activity  
JOURNAL Patent: US 5869292-A 1 09-FEB-1999;  
FEATURES  
source 1..5035  
BASE COUNT 1484 a 1127 c 1110 g 1314 t  
ORIGIN  
  
Query Match 78.7%; Score 3748; DB 9; Length 5035;  
Best Local Similarity 87.4%; Pred. No. 0;  
Matches 4374; Conservative 0; Mismatches 60; Indels 570; Gaps 2;  
  
QY 266 ccacatgcagaaatgagctccacacgtcctcctctctctgcttctgagcttctgagcttca 325  
|||||  
Db 30 CCACCATGGAATATGAGCTCTCCACCTGCTTCTTCTGCTTCTTCTGATTCGATTCGATTCG 89  
QY 326 gtgcacacagaagaatactacactggtgtcagatgagacgttcaatgagacataatgcaagt 385  
|||||  
Db 90 GTGCCACAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 149  
QY 386 atctcggtgagctgctgtgagcgaagattcctcctagagtgagcaaatcttccat 445  
|||||  
Db 150 ATCTCGGTGAGCTGCTGCTGAGCAGCAAGATTTCTCTAGAGTCCCAAAATCTTTTCAT 209  
QY 446 tcaacacactcagctgtgtacaaaagaactcgttctgtagaatcaagatgagcttca 505  
|||||  
Db 210 TCAACACCTCACTGCTGTACAAAAGACTCTGTTGTAGAAATTAACGATCACCTTTTCA 269  
QY 506 acatgcgttaagcgaagccacccctggaatggtctcgtcaggttctcattacatccagcgtgag 565  
|||||  
Db 270 ACATGCGTAAAGCCAAAGCCACCTCGATGGGTGTGATGAGTCTTACATCAGCTGAGG 329  
QY 566 ttatgatcaagtggttaacttaacacttaagaacatggttcccatcctgtcaagtctcatg 625  
|||||  
Db 330 TTTATGATACAGTGGCTCATTAACATTAAGAAACATGGCTTCCCATCCGTGCTTTCATG 389  
QY 626 ctgtgtgtgtatcctacttggaagaactctctgagagagtgatataatgagatcaagaacagtc 685  
|||||  
Db 390 CTGTGTTGTATCTTACTGGAAGCTTCTGAGGAGCTGATATGATGATGATGATGATGATG 449  
QY 686 aaagggaaagaagaatgataaagttctccctggttggaagccatataatgtctcgtgag 745  
|||||  
Db 450 AAAGGGAAGAAAGATGATTAAGTCTTCCCTGTGATGAGCATATGATGATGATGATGATG 509



QY 746 tcctgaagaagaatggtccaaatgagcctcagaccactgctccttaactacatacttc 805  
|||||  
Db 510 TCCGAAAGAGAGATGTCCTCAATGAGCTGTGACCCACTGTGCTTACTACTCATATCTTT 569  
QY 806 ctcatgtgacctggttaaaagaacttgaaatcaggccctcatgtagccctactagtatgta 865  
570 CTCATGTGAGACCTGGTAAAGACTTGAATTCAGGCTCATTTGAGCCCTACTAGATATGTA 629  
QY 866 gagaagaagagctcgtgccaagaagaagacacagaccttgacaaatttactactctttg 925  
|||||  
Db 630 GAGAAAGGAGTCTGGCCAAAGAAAGACACAGACCTTGGCAAAATTTTATACTACTTTTGG 689  
QY 926 ctgtaattgataaggaagaaaglttgagcctcagaacaagaagactccttgatgtaggata 985  
690 CTGATTTGATGAAGAGGAAAGTTGGCACTCAGAAACAAAGAACTCTTGATGACGAGATA 749  
QY 986 gggatgctgcatctgctcgtggcctgagctaaatgcaacagttcaatggttatgttaaaa 1045  
750 GGGATGCTGCATCTGCTGCTGGGCTGGCTTAAATGACACACACTCAATGGTTATGTAAACA 809  
QY 1046 ggtctcgcgaagtgctgattgagatgccaagaagaatcagttcatctattgggcatgtagtga 1105  
|||||  
Db 810 GGTCTCTGCCAGGTCTGATTTGATGATCCACAGGAAATCACTTATTTGGCATGTGATTTGGA 869  
QY 1106 tgggcaaccactcctgaaatgagcactcaataltcctcgaaggtgacacacatttctgtgagga 1165  
870 TGGGACACACTCTGAAATGAGCATCAATATCTCGAAGGTGACACATTTCTTTGTGAGGA 929  
QY 1166 acaactcgcaagcgctccttgtaaaatcctcgccaataacttctcctactcgtcctaaacact 1225  
930 ACCATCGCGACGCGCTCTGGAAATCTGCGCAATTAATCTTCTACTGCTCAAAACACTCT 989  
QY 1226 tgaaggaccttggaacagtttctactggtttgtcatatctctcccaacaaatgtagtga 1285  
990 TGATGACCTTGGACAGTTTCTACTGTTTGTGCATATCTCTTCCACACACATGTATGGGA 1049  
QY 1286 tgaagactatgtaaaatgagacagctgtccagaggaaccccaactacagaaatgaataa 1345  
1050 TGAAGGCTTATGTCAAAATAGACAGCTGTCCAGAGGAACCCCACTAGCAATGAAAAATA 1109  
QY 1346 atgaagaagcggaagactatgataatgatacttactgattctgaaatgtagtgtgtagt 1405  
1110 ATGAGAAGCGGAGACATATGATGATCTTACTGATTCGAAATGATGATGTCAGT 1169  
QY 1406 ttgattgatacactctcctccttcaatccaaatcgcctcagttgccaagaagactccta 1465  
1170 TTGATGATGACAACTCTCTCTCTTATCCAAATTCGCTCACTTGGCAAGAACATCTTA 1229  
QY 1466 aaacttggtatcattacatgtcgtcgtgaagaggaagactggtgactatgctcccttagtcc 1525  
1230 AAATCTGGGTACATTTACTCTGCTGAAGAGAGAGACTGGGACATATGCTCTTATGTC 1289  
QY 1526 tcgcccccgatgacagaagaatataaaatcattatgaaacaatggtccctcagcgatgtg 1585  
1290 TCGCCCCGATGACAGAAAGTTATAAAGTTCATATTTTGAACAAATGCGCTCAGCGATTTG 1349  
QY 1586 gtagaagaatcaaaaaaagctccgattatgagcatcacagatgaagaaccttlaagactcgtg 1645  
1350 GTAGGAAGTACAAAAAGTCCGATTTTATGGCATACACAGATGAATAACCTTTAAGACTCGTG 1409  
QY 1646 aagctatcagcatgaatcagaaatcttggaaccttacttattgagggaagtgtgagaca 1705  
1410 AAGCTATTTCACATGAATCAGGAATCTTGGGACCTTTACTTTATGGGGAAGTGGAGACA 1469  
QY 1706 cactgttgattatatttaagaatacaagcagacacataaactctaacctcaagga 1765  
1470 CACTGTGATTAATTAATTAAGATCAAGCAAGACACATATTAACCTTAACCTCAAGGAA 1529  
QY 1766 tcaactgattgctcgtccttcttattcaagaagatatacaaaaaggtgttaaaacttgaag 1825  
1530 TCACTGATGTCGTCCTCTTGTATTCAGAGAGATTACCAAAAGTGTATAAACCTTTGAAGG 1589  
QY 1826 atttccaaattctgcgaagagaataatataaaatgaatgaatgtagctgtagaagatg 1885  
|||||  
Db 1590 ATTTTCCAAATTCGCCAGAGAAATATTCAATTAATTAATGACAGTGTAGTAAGATG 1649  
QY 1886 ggcacaactaaatcaatctcgtgtgctcgtgacccgttatctactagtgttctaatatg 1945  
1650 GGCCAACTTAATCAATCTCGGTGCTGACCCGCTAATTAATCAATGTTCCCTTAATATGG 1709  
QY 1946 agagaatctagcttcaagagacatctggccctcctcactcgtcctacaaagaactctag 2005  
1710 AGAGAGATCTACTTCAGGACTCATTTGGCCCTCTCTCATCTGCTACAAATACTGTAG 1769  
QY 2006 alcaagaaggaacaagaaatgltcagaagaaggaatgltcatcctgttcttctglattg 2065  
1770 ATCAAAAGAGAAACCAAGTAATGTGACACAAGAGGAATGTCATCTCTGTTCTGATTTTG 1829  
QY 2066 atgaagaacgaagctggttaccctcagaagaatatacaagctttctcccaatccagctg 2125  
1830 ATGAGAACCGAAGCTGTGATCTCTCACAGAAATATCAACGCTTCTCCCAATCCAGCTG 1889  
QY 2126 gagtgaactggaagatccaaggttccaagcctccaacatcatgcaagaatcaatggtc 2185  
1890 GAGTCAGCTTGAAGATCCAGAGTTCCAAAGCCTCCAAATCATGACACAGATCAATGGCT 1949  
QY 2186 atgttttgatagtttgcaagttgtcagttgttgatgaggtgagcatgtagtcatc 2245  
1950 ATGTTTATATGTTATGTTGAGTTGAGTTGATGTTGTTGATGATGAGTGGCATCTGTAATTC 2009  
QY 2246 taagaatltggagacagagactgactccttctgcttcttctctctgagataatcctaaac 2305  
2010 TTAAGATTGGAGACACAGACTGACTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2069  
QY 2306 acaaatgtctatgaagaacacacactcacccatccatccatccatccatccatccatccatcc 2365  
2070 ACAAAATGCTCATTAAGAACACACACTCACCTTATTCCTTCAGAGAAATCTCTTCA 2129  
QY 2366 tgtcgtatgaaaacccaggtcattatgattctcgtggtgagcaaacactcagacttcagaca 2425  
2130 TGTGATGAGAAACCCAGAGTGTATGATTTGGGGTGCACAACTCAAGACTTGTGGAAACA 2189  
QY 2426 gaagcatgacggtccttactgaaagtttctagtgtgtgaagaagaacacgtgtgattatcag 2485  
2190 GAGGATGACCCGCTTACTGAAAGTTTCTAGTTTATTCAGAGGGGAGGAGGAGCAGC 2249  
QY 2486 aagcaagttatgaagatttca-----gaactactgctgtgaaataaca 2530  
2250 ACTATCTGACCTGGAGAAATATTCAGTGAAGACAGAGACTACATGACATGCTGCACA 2309  
QY 2531 atgcatatgaaacaaagaagcttctccagaatccaagacacctagacataggcaaac 2590  
2310 GTCTGATTGAACCAAGAAAGCTTCTCCAGAAATTCAGAACACCTAGACACTAGGCAAAAAGC 2369  
QY 2591 aatttaattggcac----- 2603  
2370 AATTTAATGCCACACAATTCAGAAATGACATAGAGAGAGACTGACCTTGTGTTGCAC 2429  
QY 2604 ----- 2603  
2430 ACAGAACACTATGCTTAATAATACAAATATGTCCTCTAGTATGTTGTGATGCTCTTGC 2489  
QY 2604 ----- 2603  
2490 GACAGAGTCTACTCCATGATGGGCTATCTTATCTGATCTCCAGAAAGCCAAATATGAGA 2549  
QY 2604 ----- 2603  
2550 CTTTTCATGATCATCATCACTGGAGCAATAGACAGTAAATACAGCCTGTCTGAATGA 2609  
QY 2604 ----- 2603  
2610 CACACTTAAGGCCACAGCTCCATCACTAGTGGGAGCATGTATTTAACCCTGAGTCAAGCC 2669  
QY 2604 ----- 2603



Db	2670	TCCAATTAGATTAAATGAGAAACTGGGGCAACTGAGATCTCTTGCTTGGGATTAAC	2729
Qy	2604	-----	2603
Db	2730	ACTATGGTACTCAGATATCCAAAAGAGTGGAAATCCCAAGAGAAAGTCCACGAAAAA	2789
Qy	2604	-----	2603
Db	2790	CAGCTTTTAAGAAAAAGATACCATTTTGTCCCTGAACGCTTGTAAGACAAATCATGCA	2849
Qy	2604	-----	2603
Db	2850	TAGCAGCAATTAAATGAGGGACAAAATAAGCCCAATAGACTCACTGGGCAAAACAG	2909
Qy	2604	-----ccacacagctcttgaaacgcatcaacggyaa	2635
Db	2910	GTAAGAGCTGAAAGGCTGTGCTTCAAAACCCACACAGTGTGAAGCGCATCAAGGGAAA	2969
Qy	2656	taactcgtactactcttcagtcagatcaagaaggyaaattgactatgatatccatatcag	2655
Db	2970	TAACTCGTACTACTCTTCACATCGAGTCAAGAGAAATTTACTATATATATACCATATCAG	3029
Qy	2696	ttgaaatgaaagaagaagatttggacattatgatgagatgaanaataaagagcccccga	2755
Db	3030	TTGAAATGAGAGAGAAATTTTGACATTTATGATGAGATGAATAATAGAGCCCCGCA	3089
Qy	2756	gcttccaaagaanaacacgacactatttatlgtcgtcagatgagaagctctggyatta	2815
Db	3090	GCTTTGAAAAGAAAACGACGACTATTTTATTTGCTGCATGTGAGAGGCTTGGGATTATG	3149
Qy	2816	ggatgtagtagtccccacatggtctctaagaacaaggtctagaagtgcagtgctcccaagt	2875
Db	3150	GGATAGATAGCTCCCCACAGTGTCTAAAGAAACAGGGCTCAGAGTGGCAGTGTCCCTCAGT	3209
Qy	2876	tcaagaagaatgtgtttcccaagaaatttaccgatgagtccttctcgaagccattatacgt	2935
Db	3210	TCAAGAAAGTGTGTTTCCAGGAATTTACTGATGAGTCCCTTTACTGAGCCCTTAATACGCTG	3269
Qy	2936	gagaactaaatgaacatttggagctctctggygcatalataaagaacagaagtltgaagata	2995
Db	3270	GAGAACTAAATGAACATTTTGGAGCTCCTGGGGCATATATAAGACAGAGTGAAGATA	3329
Qy	2996	atacatggttaacttcaagaatcaagcctctcgtcccatcttccttactatcttagctcta	3055
Db	3330	ATAATCATGTGTACTTTCAGAAATCAAGGCGCTCTGCTCCATTCCTTATTCCTAGCCCTTA	3389
Qy	3056	ttctcttctggaagaatcacaaggtcaagggacgaagacactagaanaaacttltcaagccta	3115
Db	3390	TTTCTTTGAGGGAAGATCAGAGGCAAGGACGAGCAAGAACCTTAGAAAAACTTGTCAAGCTTA	3449
Qy	3116	atgaaccocaaacttacttlttggaaggtgacaatcataatgagccaccaetaaagtagt	3175
Db	3450	ATGAAACCCAAAACCTTACTTTTGGAAAGTCAACATCATATGCGACCCACTAAAGATGAGT	3509
Qy	3176	ttgagtcacaagcttgggtactattctcgtatggttgaccttggaaaaaagtltgacaccag	3225
Db	3510	TTTGACTCCAAAGCCTGGGCTTATTTCTGTGATGTGACTGTGAAAAAAGATGTGACCTAG	3569
Qy	3236	gccgatcttgaccocctctcgtctgcacaactaacaacacttgaacccctctcatlgggagac	3295
Db	3570	GCGTGAATGGAGCCCTTGTGTGTGTGTCGACACATTAACACATGAAACCCCTCTCATGGGAGAC	3629
Qy	3296	aagtgcagtagcagaatattgctctgttlttcaacacelttgaatgagacaaagctgtgt	3355
Db	3630	AAGTGAAGATGACAGAAATTTGGCTGTCTTTTACCATCTTTGATGATGAGCCAAAAGCTGGT	3689
Qy	3356	acttcacatggaataatgagaagaagaacttgaaggtcccccgtgaataatccaatgagaatgc	3415
Db	3690	ACTTCACTGAAATATATGGAAGAACTGCAAGGGCTCCCTGCAATATCCAGATGGAAGATC	3749
Qy	3416	ccactttaaagaagattatcgcttccatgacaaatgaatggtctacataatgatatcactac	3475

D	b	3750	CCACCTTTAAAGACATTTATTCGCTTCCATGCAATCAATGGCTACATATATGATACATAC	3809
Q	y	3476	ctggcttagtaatgagccaaagatcaaaagatlcgaatgtaatcgtccaaagtgcga	3535
D	b	3810	CTGGCTTAGTAATGGCTCAGAGATCAAAAGATTGATGATATCTGCTCAAGCATGGGCACA	3869
Q	y	3536	atgaacaatccatctatctatcattcaagtgcagatglttcacgtgcgaaaaaagg	3595
D	b	3870	ATGAAACATCCATTTCTATTCAATTCAGTGCATGTGCTTCACTGACAAAAAAGAG	3929
Q	y	3596	agataaataaggacatgataacatctatccagtggttttgagacagtggaatgtaac	3655
D	b	3930	AGATAAAAATGGCACTGTACAAATCTCATTCACAGTGTTTTTGACAGTGAAGATTAC	3989
Q	y	3656	catccaagctgcgaatttcggcggtgcgaatgccttatcgagacatccaaatgctgcgga	3715
D	b	3990	CATCCAAAGCTGGAAATTTGGCGGGTGGAAATGCTTATTGGCAGCATCTACATGCTGGGA	4049
Q	y	3716	tgaacacatttcttcggtgltacagaataaagtgtcagaatcccttccttcggaatgctctg	3775
D	b	4050	TGACACACATTTTCTTGCTGTACAGCAATTAAGTGTGACATCCCTCGGAATGGCTCTG	4109
Q	y	3776	gacacatagaagatttccagaattacagcttcaggaacaataatgcagatgcgcacaaagc	3835
D	b	4110	GACACATTAGAGATTTCACATTAACAGCTTCAGGCAAAATATGGACATGGGCCCAAGAC	4169
Q	y	3836	tgagcagactcatattatccggaatcaatgaatgcctgcgagacaaagagcccttctct	3899
D	b	4170	TGGCCACACTTCATTTATTCGGATTCATATCAATGCTCGAGACACCAAGGACCTTTTCTT	4229
Q	y	3896	ggaaccaagtgaaatcgtcttgccacaaatgataltaacagcaatcaagaccagagtgcc	3955
D	b	4230	GGATCAAGTGGATCTGTGTGGCACCAATATATTTCAGGGATCAAGACCGAGGTGGCC	4289
Q	y	3956	gtcagaagcttcocagccttcacatcctcagtttaacatcatagttcagttcgtgcgga	4015
D	b	4290	GTCAAGAGTTCTCCAGCCTTCATCTCTCAGTTTATTCATCATGTATATGCTTGATGGGA	4349
Q	y	4016	agaagtgagagactatcggaggaattccacatgcagaaacctaaatgctcttcgtgcaatg	4075
D	b	4350	AGAAATGGCAGACTTATCGAGAAATTCACCTGGAACCTTAATGGCTTCTTTGGCAATG	4409
Q	y	4076	tggaatcatctcggaataaacaacaataattttaaacctcaaatatgctgcgaatcaccc	4135
D	b	4410	TGATTCATCTGGGATTAACACACAAATATTTTAACCTGCCAATTAATGTGCTGATACATCC	4469
Q	y	4136	gtttgcacccaactcatatagatctgcgaacatcttcgaatgcagttgatggctgtg	4195
D	b	4470	GTTTGGACCCCAACATTAATTAAGCATTCGAGCACCTTCTGCAATGAGTGAATGAGGCTGTG	4529
Q	y	4196	atataaagtgtgcagatgcacattgggaatggagatgaacaatatcagaatgcaga	4255
D	b	4530	ATTTAAATAGTTCAGCATCCCAATGGGAATGGAGATGAACCAATATCAGATGCACAGA	4589
Q	y	4256	ttactgcttcatccatcttaacaataatgtttgcacatgcgtctccctcaaaagtgcac	4315
D	b	4590	TTAATGCTTATCTCATTTATTAACAATATGTGTTCGCCACCTGGCTCTCTTCAAAAGCTGCAC	4649
Q	y	4316	ttcaactccaagagagagtaatgctctgcagacatcgaatgcagtgtaataatccaaagatgagc	4375
D	b	4650	TTTCACTTCACAAAGGAGAGTAATGCTCGAGACCTCAGGTGAATATCCAAAAGAGTGGC	4709
Q	y	4376	tgcaagtggaacttcggaagacaaatgaagatcaagagataactatcaaggagataaat	4435
D	b	4710	TGCAAGTGACTTCACAAACACATGAAGATCACAGAGATAACTCATCTCAGGAGATTAAT	4769
Q	y	4436	ctcgcgtaccagaatgatgtgaagagttccatctccagcagatcaagaatgcac	4495
D	b	4770	CTCTGCTTACACAGATATATGTAAGAGATTCTCTCATCTCCACAGCTGAAGATGGCCATC	4829
Q	y	4496	agtgaaatcctcttttccagaatgcgaagataaagttttcaaggaaatcaagactcct	4555
D	b	4830	AGTGAATCTCTTTTTCAGAAATGGCAAGTAAAGTTTTCAGGGAATCAAGATCTCT	4889



QY	4556	tcacacctctgtgtgaactctctcttagacccaacggttactgactgcgtacactctgaaatcacc	4615
Db	4890	TCACACCTGTGTGTAACCTCTTAGACCCACCGTTACTGACTGGCTAACCTTCGAATTCACCC	4949
QY	4616	cccaagatctgggtgcacacgaattggcccttaagatggagttctctggagctgcagagcagag	4675
Db	4950	CCCAGATTGGGTGCACCGATTGGCCCTTAGAGATGGAGGTTCTGGGCTCGAGGCAAGG	5009
QY	4676	acctctactgactcgagcgagtc	4699
Db	5010	ACCTCTACTGAGGTTGGGCACTGC	5033
RESULT	8		
LOCUS	AR071306	5035 bp	DNA
DEFINITION	Sequence 1 from patent US 5910481.		PAT
ACCESSION	AR071306		
VERSION	AR071306.1	GI:7222194	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 5035)		
AUTHORS	Voorberg,J.J.		
TITLE	Hydrid proteins with modified activity		
JOURNAL	Patent: US 5910481-A 1 08-JUN-1999;		
FEATURES	Location/Qualifiers		
source	1..5035		
BASE COUNT	1484 a 1127 c 1110 g 1314 t		
ORIGIN			
Query Match	78.7%;	Score 3748;	DB 9; Length 5035;
Best Local Similarity	87.4%;	Pred. No. 0;	
Matches 4374; Conservative	0;	Mismatches 60;	Indels 570; Gaps 2.
QY	266	ccacatgtcaaatagagagctctccacactgctctctctgtgcctttggagatcttgctta	325
Db	30	CCACCATGGAATAGAGCTCTCCACCTGCTTCTGTGCGCTTGGCGATTCTGCTTTPA	89
QY	326	gtgcacacgaagaatactacacccctgggtgctgagactgtcaggaactatgcaaaagt	385
Db	90	GTGCCACCGAAGATPACTCTGGGGGACGTGGAACTGTCATGGGACTATATGCAAAAGTG	149
QY	386	atctcggtagctgctctgtgcagcaagaattctcctctagagtgccaaatcttccat	445
Db	150	ATCTGGTGAGTGGCTGTGGAGCGAAGATTTCCTCTCTAGAGTGGCCAAATCTTTCCAT	209
QY	446	tcaaacactcagtcgtgtacaaaagaactcgttctgtagaatcaacggttcaaccttca	505
Db	210	TCACACCTCTCACTCGTGTCAAAAAGACCTCTGTTGTAAATTCACGGATTCACCTTTTCA	269
QY	506	acatcgctaaagcaagggccacacctgagatggtcgtctagtgctctacatccagagctgag	565
Db	270	ACATCGCTAAGCCAAAGGCCACCTCGATGGGTCTGCTAAGTCTTACATCCAGGCTGAGG	329
QY	566	tttaatgatacagtggtlcatlatacactlaaagaacatggtctccacactcgtcagttcttca	625
Db	330	TTTTATGATTCAGTGGTCAATTACACTTAAGACATGGCTTCCCATCCTGTCAGTCTTCATG	389
QY	626	ctgtgtgtgtatcctcactcgtgaaagctctcgagggagctgaaataatgatgatcagaaccagtc	685
Db	390	CGTGTGTGTATCCCACTGGAAGAGCTTCTGAGGGAGCTGAATATGATGTCAGACAGTC	449
QY	686	aaagggaagaagaagatgataaagtcctccctcgggtggagagccatacatatgtcctggcag	745
Db	450	AAAGGAGAAAAGATGATTAAAGTCTTCCCTGGTGGAGGCCATACATATGTCTGGCAGG	509
QY	746	tctctgaaagaaatgtgtccaatgctccttgacccaactgtgcttcaactaataatctt	805

[illegible]



QY	1886	ggccaactaaatcaagatcctcggctgcctcgaacccgcgtatctactcgttgcgttaatatg	1945
Db	1650	ggccaaactaaatcagatcgaatcccggtgcctcgaacccgcgtattactcgtattgaatgg	1709
QY	1946	agagaagctcagcttcaagacatcttggccctctctcactcgcgaagaatcttag	2005
Db	1710	agaagatcttagcttcagacatcttggccctctctcactcgcgaagaatcttag	1769
QY	2006	atcaaaagggaacaccagataatctgcagacaagaaggatctaccctcgtttctglatly	2065
Db	1770	atcaaaagggaacaccagatattgctcagacagaagaaatgcatcctgthttctgaaattg	1829
QY	2066	atgagaaaccgaagctcgttactctacagagaataatacaagctttctcccaatccagct	2125
Db	1830	atggaacccgaagctggtactctacagagaataatcacaagctttctcccaatccagctg	1889
QY	2126	gagtcagctctgagatcccaagcttccaagcttccaacatctgcacagaatcaatgct	2185
Db	1890	gagtcagctctgagatcccaagcttccaagcttccaacatctgcacagaatcaatgct	1949
QY	2186	atgtttcttgatagtttgcagcttgctcaagtttgttgcaagagtgtagaactcgttacatc	2245
Db	1950	atgttttttgatagtttgcagcttgccttcagtttggttgcatgagtgtagaactcgttacatc	2009
QY	2246	taagcatctggaagacaagacagacttccctctgcctctctctcgtgatatccctcaac	2305
Db	2010	taagcatctggaagacaagacagacttccctctgcctctctctcgtgatatccctcaac	2069
QY	2306	acaaaatggtctatgaaagacacatcccatctccatctcagagaagaactgcttca	2365
Db	2070	acaaaatggtctatgaaagacacatcccatctccatctcagagaagaactgcttca	2129
QY	2366	tgctgataggaaaacccaggtctatgattctctgggtgcacaactccagactttgggaca	2425
Db	2130	tgctgataggaaaacccaggtctatgattctctgggtgcacaactccagactttgggaca	2189
QY	2426	gaggaagacagcgccttactcgaagtttctcagttgtgcagaagaacacbtggtatattacg	2485
Db	2190	gaggaagacagcgccttactcgaagtttctcagttgtgtgattgcagaagggagagagacacg	2249
QY	2486	aggaagctatgaaagatatcca-----gcatactctgcagtaaaaaa	2530
Db	2250	actatctgcgacctgggaagaatatttcagtaagaagacagacatccacatctcgtccaca	2309
QY	2531	atgcacatcgacaagaagaagctctcccaagatctcaagaacacccctgacagcagaaagc	2590
Db	2310	gtctgattggaaccaaagaagcttctcccaaatctcaagacacccctgacagcagaaagc	2369
QY	2551	aatttaatgcac-----aatttaatgcac-----	2603
Db	2370	aatttaatgcacaccaaatttccagaaaatgacatgagaagacatgaccccttgattgcac	2429
QY	2604	-----	2603
Db	2430	acgaacaaactatgcttaaaatgcaaaatgctcctctgattgatttggtaatgctcttgc	2489
QY	2604	-----	2603
Db	2490	gacagagctctactccacatgagggctattccttattgattctccagaagacaaatattgaga	2549
QY	2604	-----	2603
Db	2550	ctttttctgattgattcattcactcgtgagcaatgagacatgaatgaacagcctgctgaaatga	2609
QY	2604	-----	2603
Db	2610	cacacttcagggcacaagctccatcacaagtgggagacatgattttacccctgaagcagacc	2669
QY	2604	-----	2603
Db	2670	tccaatattgaattaaattgagaaactggggacaaactgcagatccttcttgctgggataacc	2729

[illegible]







```

Db 543 CACGTGACCGTGGTGAAGACCTGAATTCGGGCTCATTTGGAGCCCTGGTGGTTGTGAGA 602
Qy 868 gaagagagcttgcccaagaagaacacagacctgcacaaatttactacttctgtc 927
Db 603 GAAGGGAGTCTGACAGCAAGAGAACCTCGACCAATTTTGACACTTTTGTGCT 662
Qy 928 gtattatgaaggaaatttgccactcagaaacaaagacctctgtatgacagatag 987
Db 663 GTCTTTATGAGGAGAAAGTTGGACTCTACAGCAAGAAATGCTCTTGACAGGGGCATG 722
Qy 988 gatgctgacatgctgcggcgctgccttaaaatgacacagtcacatgttactgtgaag 1047
Db 723 GATCCCACTCTGCGACGAGGCCCAACCTGCAATGACACAGTCAATGGCTATGTCAACAGG 782
Qy 1048 tctctgcagagcttgatgtatgtacacagaagaatcagcttaattggaatgtgaa 1107
Db 783 TCTCTGCCAGGTCTGATCGGATGCATGAAGAAATCACTTCACTGTGCAAGTATGGAAATG 842
Qy 1108 ggcacacactctgaagtgcactcaatattcctggaagtgacacattctctgtgaag 1167
Db 843 GGCACACGCCGGAGAGTGCATCTCATTTTCTTGAAGGCCACACGTTTCTGTAAGGCAC 902
Qy 1168 caccgcagagcgctccttggaatctgcacaataacttccttacttactgcacaacct 1227
Db 903 CATGCCAGGCTTCTTGAGATCTGCGCACTAACTTCTCTCACTGCTCAGACATTTCTG 962
Qy 1228 atggaacttggaagttctactgtttgttcatactctctccacaacacatgtatg 1287
Db 963 ATGACACTTGGGACTCTCACTGTTTGTTCATATCTCTTCCACACCCACATGGTGCGCATG 1022
Qy 1288 gaagctatgtcaaaatgagcagctgtccacagagaaccccaactacgaatgaaataat 1347
Db 1023 GAGGCTACCTCAGAGTAGAAGCTGCGCGAGAGGCCCAAGCTGCGGAGAAAGCTGAT 1082
Qy 1348 gaagaaagcgaagactatgtatgtacttactgtatctgaaatgtatgtatgtat 1407
Db 1083 GAAGA---GGAAGATATATGATGACAAATTTGTAGACATCGGACATGACATGGTCCGCTC 1139
Qy 1408 gatgatgacaactctccttcccttataccaatccgctcagttgcacaaagaactc 1467
Db 1140 GATGGTACACAGCGTGTCCCTTATCCCAATCCGCTGGTTCGCAAGAGCATCCCAAA 1199
Qy 1468 acttggtatcatcatctgtctgcaagagaaggaactgggaactatgtcccttact 1527
Db 1200 ACCTGGGTGACATCACTCTGCGACAGAGAGAGGACTGGGACTTAAGCCCCGGGCTCCC 1259
Qy 1528 gcccccgatgacagaagttaaaagtcaatatttgacaacatgcccctcagcgat 1587
Db 1260 AGCCCGATGACAGAAAGTATTAAGTCTCTACTTGAAAGTGCTCTCAGCGAATTGGT 1319
Qy 1588 aggaagttaaaaaaagctcgatattatgtacatacagatgaacacttgaagctgt 1647
Db 1320 AGGAAATACAAAAAGCTCATTTCTGCTGCTTACACGATATTAACATTAAAGACTCGTAAA 1379
Qy 1648 gctatcagatgaatcagaatcttggaacttacttattatgaggaagtgtgagaaca 1707
Db 1380 GCTATTCGATGAATCAGGAATCTGAGACTTTTACTTTATGGAGAGTTGGAGACACA 1439
Qy 1708 ctgttgatlatattaagaatcaagaagaacatataacatctaccctcagatc 1767
Db 1440 CTTTGTATTAATTAAGATTAAGAGCGAGCGACATATTAACATCTTACCTCATGGAATC 1499
Qy 1768 actgatgtccgtccttgtatctcaagagatcaaccaagggtgaataacttgaag 1827
Db 1500 ACTGATCTTCAGCGCTTGGACCCAGGAGGACTTCTAAAGGTTGGAATTTGAAAGAC 1559
Qy 1828 ttccaacttgcacagagaataatcaaatataaagtgcagctgacatgcagaagatg 1887
Db 1560 ATGCGCAATTTCTGCCAGAGAGACTTTTAACTATTAATGAGACAGTGTGAGAGATGGG 1619
Qy 1888 ccaactaatcagatccctggtgctgacccgctatcactagtttcgttaatatgag 1947

```

---

```

Db 1620 CCAACCAAGTCCGATCTCGTGCCCTGACCCGCTACTACTCGAGCTCATTAATCTAGAG 1679
Qy 1948 agagatcagcttcaggaactcattgcccctctcctcactctgtcacaagaatctgtat 2007
Db 1680 AAAGATCTGCTTCGGGACATCATTTGGCCCTCTCTCATCTGTCAACAAGATCTGTAGAC 1739
Qy 2008 caaagagaacacagataatgtcagacaagaagaatgtcaatccgtttctgtattgtat 2067
Db 1740 CMAAGAGAAACCAAGATGATGTCAAGACAGAAAGCAAGCTCATCTCTTTCTGTATTTGAT 1799
Qy 2068 gagaacccaagctgtgtaacctcacaagaatalatacaagcttctccccaatccaagt 2127
Db 1800 GAGATCAAAAGCTGTGATCTCGCAGAGATATTTCAGGCTTCTCCCAATCCGAGATGGA 1859
Qy 2128 gtgcagcttgaagatccaaggttccaagctccaacatctgacacagatcaatgtat 2187
Db 1860 TTACAGCCCGAGATCCAGAGTTCCAAAGCTTCAACATCATATGACACACATCAATGGCTAT 1919
Qy 2188 gtcttgatagtgtgacagttgtcagttgtgttgatgaggtgacatctgtacattct 2247
Db 1920 GTTTTGTATAGCTTTCGACAGCTGCGGTTGTTTTCACAGAGGTGCGATATCTGTACATTCTA 1979
Qy 2248 agcatlgaagcacagactgacttcttctgtcttcttcttctgtatataccttcaaac 2307
Db 1980 AGTGTGAGCACAGACAGGACTTCTCTCGCTCTCTCTCTCTGCTGCTACCTTCAAAACAC 2039
Qy 2308 aaaaatgtctatgaagacacactcaaccctatcccaatctcagaagaactgttcat 2367
Db 2040 AAATGCTCTATGAAGACACACATCACCTCTTCCCTTCCAGAGAAAGCTTCTCATG 2099
Qy 2368 tcatgtgaaacccaagttcatgtatctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2427
Db 2100 TCAATGGAAGAAACCCAGGCTCTGAGTCTAGGGGCCACCACTGAGCTTGGCGAAGACA 2159
Qy 2428 ggcatagcgccttactgaaggttctcagttgtgtgacaagaacacatgtgtatctac 2487
Db 2160 GCGATGACACCTTACTGAAGGTGTATGTATGTGACAGGAGATGTGTATTAATGAT 2219
Qy 2488 gacagttatgaagatattcagcatctgtcgtatgtatgaataaacaatgtcattga 2547
Db 2220 AACACTATATGAATATATTCAGAGCTTCTTCTGCTGAGTGAAGAAATGATCAATGAAC 2279
Qy 2548 agcttctccagaattcaagaacacctagcactaggaagaagaatttaatgtccac 2607
Db 2280 ----- 2279
Qy 2608 ccagcttgaaacgcacacacggaataactcgttactactcttcaagatcaagag 2667
Db 2280 -----GACATTAAGCTTCTCTTCTTCTGAGCGGAGAGAC 2315
Qy 2668 gaaatgacatgatgtatccatactcagttgaaatgaagaagaattgtgacattat 2727
Db 2316 AAATGAGCATATGATGATATCTTCTCAACTGAAGACGAAGGAGAAAGATTTTGACATTAC 2375
Qy 2728 gatgatgataaatacgaagcccccagcttccaagaagaacacacacacattat 2787
Db 2376 GTGAGGATTAATATCAGGACCTCGCAGCTTTCAGAGAGAAACCGACACTATTTTCATT 2435
Qy 2788 gctgcagttgagagagctcttggaattatgagatgtagtcctccacaatgttctaaga 2847
Db 2436 GCTGCGGTGAGACAGCTCTGAGATTAAGGATGAGCAATCCCGCGGCGCTAAGAAAC 2495
Qy 2848 agggctcagaagtgtgcagttgcctcagttcaagaagaatgttctccaagaattac 2907
Db 2496 AGGCTCTCAGAACGAGAGGCTCGGTTCAAAAGGTGTCTTCCGGGAATTTGCTGAC 2555
Qy 2908 ggtctcttactaagcccttaacacgtgtgagaactaaatgaacatttggagaccc 2967
Db 2556 GGTCTCTTCACGACGCCGTGTACCGCGGGAACCTCAACAACATTTGGGCTCTTGGGA 2615
Qy 2968 ccatataaagacagaagtltgaagataatatacatgttaacttccagaaatcagcct 3027
Db 2616 CCTACATCAGACGGAAAGTTGAAGACAAACATCATGTATCTTCAAAAACAGGCGCT 2675

```



QY	3028	cgctccattcttcctcatcttagcctatttcttaaggaagatcaagagcaagagca	308
Db	2676	CGTCCCTAATTCCTTCTACTCGAGCGTTTATTTCTTATCCGATGATAGGAGAAAGGGCA	2735
QY	3088	gaacctgaanaaaccttctgaagccctaatgaaaccaaacacttaactttggaagtgcaa	3147
Db	2736	GAACCTGCACACAACTTCGTCTCAAGCCAAATGAACCAAGAACTTACTTTTGGAAAGTCAG	2799
QY	3148	catcatatgacccactaaagatgatttgaactgcaagccctgagcttatcttcctgat	3207
Db	2796	CATCACTATGGACCCACAGAAAGAGATTTGATCTGCAGAAAGCCCTGACTTTTCTGAT	2855
QY	3208	gttgactctgaaaaaagttgtgcaactaagcgccttatgtgaaccccttcctgtgcacaact	3267
Db	2856	GTTTACCTCGGAAAAAGATGTGCATCGCTGATGAGGCCCTTCTGTGATCTGCGCGCC	2915
QY	3268	aacacactgaacccctgtgtcaatggagaagaatgacaaagaaatgttcctgtcttcttc	3327
Db	2916	AACACCTGAACGCTGCTCAACGGTAGACAAAGTACCGTGCAGAAATTTCTCTGTTTTTC	2975
QY	3328	accactcttgaatgagaccaaaagcctgttacttcaactgaaataatggaagaaactcgag	3387
Db	2976	ACTTTTGTGATGAGACAAAGACGTGTTACTTCTACGTGAAAAATGTGGAAAGCAACTGCCG	3035
QY	3388	gtctcccgcaaatccgaagtggaaagatcccaactttaagaagaatattatgtcttcatga	3447
Db	3036	GCCCCCTGCCACTCGCAGATGGAGAGACCCCACTCTAAGAAAGAACTATGCTTCCATCA	3095
QY	3448	atcaatggtctacataatggaatacaactcctgtccttagtaatggtctaaagatcaaaagat	3507
Db	3096	ATCAATGGCTATGTGATGATACACTCCCTGGCTTATGTATGGCTCAGATCAAGATGC	3155
QY	3508	cgatggtatcttgcctcagcataggcgcaagcaatgaaacaatccattctatcttaagtga	3567
Db	3156	CGATGATATCTGCTCACACATGGGCGAGCAATGAAAAATTCATTCGATTAATTTTGGCGA	3215
QY	3568	catgtcttcaactgtgacgaaaaaaggaggtataaagtgcgaactgacaactctcatca	3627
Db	3216	CACGTGTTCACTGATGACGAAAAAGAGAGATATAAATGGCCGTGTACATCTTATACG	3275
QY	3628	gggtcttcttgagacagltggaatgltaccatccaaagctgaaatttggcgggtggaatgc	3687
Db	3276	GGTGTCTTTTGAGACAGTGGAAATGCTAACCTCCAAAGTTGGAATTTGGCGATTAAGAACG	3335
QY	3688	cttatgtggagacatctacatgcttggatggagaacacttttctggttgaaagcaataag	3747
Db	3336	CTGATTTGGCGAGCACTCGAAGGTGGGATGAGCAGCACTTTCCTGGTATCAGCAAGAG	3395
QY	3748	tgtaagaactccccggaatgagctctctggaacacatataagatttcaagtttcaactctca	3807
Db	3396	TGTAGGCTCCACTGAGGAATAGGCTCTTCGAGCAGCATTTAGAGATTTTCAATCACACTTCA	3455
QY	3808	ggacacataatggaacgttggccccaagacttggccaagatctcatatctccgatcaatcaat	3867
Db	3456	GGACAGATATGAGCAGTGGGCCCAAGCTGGCCAGACTTCATTAATTCGGATCAATTCAT	3515
QY	3868	ggcttggagacccaaggagcccttctcttgaaacaagttgatctgttggcaccacaatgatt	3927
Db	3516	GCTTGGAGCACAGAGATCCCACTCTCTGATTAAGTGAATCTGTTGGCACCAATGTATC	3575
QY	3928	attcaaggaatcaagaaccaagggttgcccgltcagaagtlttccaagctctacatctctga	3987
Db	3576	ATTTCAGGCAATCATGACCAGGGGTGCCCGCAGCAAAATTTTCCAGCCTTACATCTCCAG	3635
QY	3988	tttatcatcatgtatagttcttgaatgggaagaagtgtgcagacttalcgaggaatctcaact	4047
Db	3636	TTTATTCATCATATGATCACTTTTGAAGGGAGAACTGGCAAGATTTACCGAGGAATTTCCACG	3695
QY	4048	ggaaccttaatgtctctcttggcaatgtgatatctcgggataaaacaatatctttt	4107
Db	3696	GGCACCTTAATGGTCTTCTTTGGCAATGTGACGCAATCTGGGATTAACACATATTTT	3755

QY	4108	aaocccaatattgctgcgataacacgcgtttgaccccaactcaattatagattccgagc	4167
Db	3736	AACCTTCGATGTGCGCTCGGTACTCCGTTTGCAACCAACATTACAGCATCCGCGAC	3815
QY	4168	actctgcgatgagttbatggtcgttgattaaatagltgcagatgcatttggaatg	4227
Db	3816	ACTCTTGCATGAGATTGATGGCTGTGATTTAAACATTTCAGCATCCCTCGGAATG	3875
QY	4228	gagagtaagaacataatcagatgcacagatactatgcttcatcctacttaccataatgtt	4287
Db	3876	CAGATTAAGGATATCGATCGACATCAAGATACAGGCGCTCTCCCACTAAGCAATATATT	3935
QY	4288	gcaacctggtctccttcaaaagcttcgacttcaactcctcaaggagagatgatgcttgaga	4347
Db	3936	GCCACCTGCTGCTCTTTCACAAAGCCGACTTCACCTCCAGGGGCGGACGAATGCTTGCGGA	3995
QY	4348	cctcaggtggaataatccaaagaatggtctgcgaatgagatgcacgaagaacatgaagt	4407
Db	3996	CCCCGGGTGAGCAGCGCGCAGAGAGATGGCTGCAGAGTGGACCTGCAGAAAGACGGTGAAGGTC	4055
QY	4408	acaaggaagtaactactcaaggagtaaaatctctgtcttcaacagatgtaatgtaagagttc	4467
Db	4056	ACAGGCATCACACACCCAGCGGCGTGAAGTCCCTGCTCAGCAGCATGTATGTGAAGAGTTC	4115
QY	4468	ctcaatctcagcagatgcagaatggtccatgaatgagatctcttlttcgaatggtcgaagta	4527
Db	4116	CTCGGTGCAGATGACAGGACGGCGCGGTGGACCCGTGTTCTTCAGAGAGGGCACACG	4175
QY	4528	aaggttttcagggaaatccaagactccttcaacctgtggtgtaactctctagaccacag	4587
Db	4176	AAGGTTTTCAAGGGCAATCAGAGACTCTCTCCACCCCGTGTGAGAACGCTCTGGACCCCG	4235
QY	4588	ttactgactgcctaccttgcgaattcaaccaccagatgtggtgtgcacagatgtccctgag	4647
Db	4236	CTGTTCACGCGCTACCTGAGATCCACCCCAAGACTGGCGCAGACATGCGCCTGAGG	4295
QY	4648	atggaaggtctctggtcgtgcgagcaaaagactctactga	4686
Db	4296	CTCGAGGTTCTTAGATGTGAGGCAACAGGATCTCTACTGA	4334
RESULT	10		
LOCUS	ARI10041	4451 bp	DNA
DEFINITION	Sequence 42 from patent US 6114148.		
ACCESSION	ARI10041		
VERSION	ARI10041.1	GI:12826317	
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 4451)		
AUTHORS	Seed, B. and Haas, J.		
TITLE	High level expression of proteins		
JOURNAL	Patent: US 6114148-A 42 05-SEP-2000;		
FEATURES	Location/Qualifiers		
source	1..4451		
BASE COUNT	972 a 1568 c 1234 g 677 t		
ORIGIN			
Query Match	54.7%; Score 2605.4; DB 9; Length 4451;		
Best Local Similarity	74.8%; Pred. No. 0;		
Matches 3266; Conservative 0; Mismatches 1101; Indels 0; Gaps 0;			
QY	328	gccacccagaagatactactcctgtgtgcagtggaactgtcatgagactatatagcagaatgtag	387
Db	84	GCCACCCGCCCTACTACTCTGCGGCCGTGGAGCTGTCCTGCGGACTCATCATCAAGAGCGAC	143
QY	388	cttggttagctgcctctgtagcagcaagaatttctctctagtagagtgcaaaatctttccatc	447
Db	144	CTTGAGCGAGCTCCCGTGTGACGCGCGCTTCCCGCCCGCGTGGCCCAAGACTTCCCTTC	203



OY	448	aacaccatcagtcggtatcaaaaagactctgtttgtgaattcaagtggttaccctttcaac	507
Db	204	AACACCAACCGTGGTGTCAAGAAAACCCCTGTTGTGTGAATTTTCAACGACACCACTGTTTAAAC	263
OY	508	atcgctaaagccaagaagccaacctgtagtggctctgagctcagctccacaacccaggtcagatt	567
Db	264	ATTGCCAAAGCCGCCGCCCTCTGGATGCGCTCTTGCGGCCACCACTATCCAGGCCAAGGTG	323
OY	568	taagatacaagtgtgaltactactactaaagacatggtcttccatctctgacgttctcatgct	627
Db	324	TACGACACCCGGTGTGATCACCTCGAAGAACATGGCCAGCACCCCGTGCAGCTCGACAGCC	383
OY	628	gttggtgtatocactactctggaagactctctggagagcgaatatgatgatgacccaagtaa	687
Db	384	GTGGCGGTGAGCTACTGTGAAGGCCAGCGAGGGGGCCGAGTACGACGACCAAGCTGCCAG	443
OY	688	aggagagaagaagaatgtgataaagcttcctccctggtgtagaagcatataatgtctgcgaagtc	747
Db	444	CGCGAGAAGGAGGAGCGACCAAGGTGTTCGGGGGGGAGCCACACCTACGTGTGAGAGGTG	503
OY	748	ctgaaagaagaatgtgtccaatgagcctctgacccaactgtgcttactactatatactttct	807
Db	504	CTTAGAGGAGAACGGCCCTATGGCCACGACGCCCTCTGCTCCTACCTACAGTACTCTGAGC	563
OY	808	catgtgaccctggttaaaaagacttgaattcaagcctcatgtggacccactagtatgtaga	867
Db	564	CACGTGACCTGGTGTGAAGGATCTGMAACAGCGGGCTATGGCGCCCTGCTGGTGTGTGC	623
OY	868	gaagagagctctgyccaagagaaaagacacagactctgacaaattatatactacttttct	927
Db	624	GAGGGCAGCGCTGGCCCAAGGAGAAACCCAGACCCCTGCACAACTTATCTCTGCTGTTCGCC	683
OY	928	gtatttbataaaggaanaattgtgcactgaagacaanaaactcttgatgtgagatag	987
Db	684	GTGTTTCGACGGGGGAAGACTGTGGCACAGCGAAGCTTAACAAAGCCTGTGATGACGAGACCG	743
OY	988	gattgtcatctgtctcgggctctggcctlaaattgcacacagctaaatggtatgtlaaacag	1047
Db	744	GAGCCGCCGACGGCCCGCGCTGGCCCAAGATGCACACGTTAACGGCTAGCTGAACGCG	803
OY	1048	tctctgcgaagttctgatatgtgataccacaggaataactgctatattgtaagtgtgaaatg	1107
Db	804	AGCTTGCGCCCGCTGATCGGCTGCTCCACCCGAAAGGGTGTACTGTGACAGCTCATCCGCAATG	863
OY	1108	ggaacacatcctgtagatgacactcaatatctcctggaagtcacacattctctgtgagagac	1167
Db	864	GGCACACACCCCTAGAGTGGCACACAGATCTTCTGTGAGAGGCCACACCTTCTGTGGCGAAC	923
OY	1168	catgcgcagcgctcctgtgaatatctgcgcaataacttttcttactctgtctloaaacactltg	1227
Db	924	CACCGCCAGGCGACGCGTGGAGATCACCCCATCACCTTCTGACAGCCACGACCTGCTG	983
OY	1228	atggaacttggagaatttctaactggttttttgcataatctctccacaacaatgatgagag	1287
Db	984	ATGGAACCTAGGCCATGTTCTGCTGTTCTGTGCACATATAGCAGCCACCAAGACAGCGGGCATG	1043
OY	1288	gaagcttatltcaaaagttagacaagtgtlccagagaagaaacccaactacagatgaaanaatat	1347
Db	1044	GAGGCTTACGTGAAGGTTGGACAGCTGCCCGGAGGAGCCCCAGCTGGCATGAAAGAACAC	1103
OY	1348	gaagaagacggaagactatgtatgtatgtacttactgtatcttgcgaatgtagtgtgcaagttt	1407
Db	1104	GAGGAGGCCGAGGACTTCAGCAGCAGCAAGCTGTACCGAAGGAAATGGAATGTCTATGAGCTTC	1163
OY	1408	gatatatcaaacctctcttcttacttcaaatgcgtcagttggttgcacaagaagatcataa	1467
Db	1164	GAGCGACGACAAAGACCCCACTTATTCAGAGTCCGAGGCTGGCCAAAGAGACCCCTAAG	1223
OY	1468	acttggtgataactatcatctgtctgtcgaagagagagactggtgactatgtctccctttagctc	1527
Db	1224	ACCTGGGTGACATCATTCGCCCGCGAGGAGGAGATGTGGATCTACGCCCCGCTATTACTTG	1283

QY	1528	gccccgatgacgaagttataaagtcaatatttgaacaatvgccctcagygatttgt	1587
Db	1284	gcccccgacgcgcccacacttccaagagccagctacttcaaacgagcccgacgacattccgc	1343
QY	1588	aggaagtacaaaanaagtcgcatattatggcacaacagatggaaccttaagaactcgtyaa	1647
Db	1344	gcgacgtrtcaagaaagctgccccttcatttgccttaccacgcagacgacttttcaagaccgcgag	1403
QY	1648	gctatcagcatgaatcaagaaatcttgggaaccttcttatgggggaagtggagacaa	1707
Db	1404	gcgatccagacgcagatccggcacttcttgcggccctctgcttaccgggagagtgggcgacac	1463
QY	1708	ctgttgattatatttaagaatcaagcaagcagacataacatacctcaegaaac	1767
Db	1464	ctgctgtatcatctttcaagaacacagggcagagcccttcaaacatcttacccccgcgcatct	1523
QY	1768	acgatctcgcttcctttgttatccaagagatcaccaaaagtgttaaacatttgaagat	1827
Db	1524	accgacgtggcccccctgtacagccggccctgcaccaaggggtgagacacttgaaagac	1583
QY	1828	tttccaattctgcagaggaagaataatcaataaayggacagtgactgtagaagatvg	1887
Db	1584	ttmcccatctcgtggccggcgagattcttcaatcaaatgtgacacgtgacgtagagagcgc	1643
QY	1888	ccaactaatcagatcctctggtygccttgacccgctatctactagtttcgttaatatggag	1947
Db	1644	cccacccaagacgcagccccccctgcctgacccctactacagcagcttgcgtgaacatvgag	1703
QY	1948	agagatctagcttagagaccatcttgccctctctctcatctgcttaanaagaatctgtagat	2007
Db	1704	cgccacactcgccctccggacgtatctggcccccctgcctgatatctgtacaaagagacgttgac	1763
QY	2008	caaaagagaacacgaataatgctcaagcaagaaggaatgctatccttcttctgattgat	2067
Db	1764	cagggcgcgacacacgactatcatgagacgacmaagcgaacgtaactctgttcagcgtgtttcac	1823
QY	2068	gagaaacgaagctgtgtaacctcaagagaaatatacaacgcttctccccaatccagctgga	2127
Db	1824	gagaaacgagcagctggtatctgaccggaamaactcaacgcttccctcccaaccccgctggc	1883
QY	2128	gtgcagcttgagatcccaaggttccaaagctccaaatatatgcaagatcaatgtgcat	2187
Db	1884	gtgcagcctggaagatcccgaggttccagggccaggaamaactatccacagactcaacggctaac	1943
QY	2188	gttcttgatagtttgcaagttgctcaagttgttgcaaggtggcaactcagctacatcta	2247
Db	1944	gtgtttccagacacctgcagactgagacgtgtgccttgcatagagagtggtgacctgactgacatctctg	2003
QY	2248	agcattgagagacagactgtaacctcttctgctcttctctctctgtaataccttcaaacac	2307
Db	2004	agcatvgcgcgcccgacgcgacttctgagaggtgttcttccggggatvatccattcaagcac	2063
QY	2308	aaaaatgtctattgaaagcacacatccacctattcccatcttcaagsgaaactgltctatg	2367
Db	2064	aaagtgtgtgaacgagacacacctgacacctgttcccttctccgggagaaactgtgttcaatg	2123
QY	2368	tcgaatgaaaacccaagttctatgattcttggggttggcaaacactcaagacttccogaaacga	2427
Db	2124	ttcttggagaaaccccgccctgttgatattcttgggctgcacaaacagcgaacttccgacacggc	2183
QY	2428	ggaatgacgccttactgaaaggttcttgaatttgtaaaaagaacactggtgatattacgag	2487
Db	2184	ggcatatgactgcgctgctgaagatctccagctggagaaamaacacggggagactgactgacag	2243
QY	2488	gacagttatagaagatatcttaagcatcttgcgtgagttaaaacaatgccaatggaaccaaga	2547
Db	2244	gacagctacagagacacatctccgctacgcttgcacaaamaacacgcatgagaccccgcc	2303
QY	2548	agctctcccgagaatcaagaacccctagcactaggcagaagaacttaatggcaccccca	2607
Db	2304	ttccttttcccaaaactcccgccacaccccgacgacgctgacagacacttttaacgccaacccc	2363
QY	2608	ccagttcttgaaacgcacatacagggaaataaactcgtactaactcttcagtcagataaagag	2667











[illegible]

RESULT	12	
AR029065		
LOCUS	AR029065	9009 bp
DEFINITION	Sequence 1 from patent US 5859204.	DNA
		PAT
		29-SEP-1999

ACCESSION	AR029065	
VERSION	AR029065.1	GI:5941038
KEYWORDS		
SOURCE	Unknown.	

ORGANISM	Unknown.
REFERENCE	Unclassified.
1	(bases 1 to 9009)

AUTHORS	Lollar, J. S.
TITLE	Modified factor VIII
JOURNAL	Patent: US 5859204-A 1 12-JAN-1999;
FEATURES	Location/Qualifiers
source	1. 9009

	/organism="unknown"
BASE COUNT	2853 a 1907 c 1844 g 2405 t
ORIGIN	

Query Match	49.0%;	Score 2336.4;	DB 9;	Length 9009;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2337: Conservative	0;	Mismatches	1;	Indels 0;
				Gaps 0;

[illegible]

D	390	CGATTAGCCAAAGGCCACCCCTGATGGGGTCTGCTAGTGGCTTACCATCCAGCGAGAGTTTAA	443
Q	570	tgatacagttgltcatlaacacttaagaacatlggtctccacatcgtlcatgctcatalgctgtc	629
D	450	TGATTCACGTGGTCTCATTTACACTTAAGAACATGAGGCTCCCATCTGTCAGTTCATCATGTCT	509
Q	630	tgtgtatccctacacggaagctctcttgaggagcttgaaatgatgatcgaacagtcgaag	689
D	510	TGTGTATCTTACCTACCTGGAAGCTTCTTGAGGAGCTGTGAATATGATGATCAACACAGTCAAG	569
Q	690	ggaagaagaagaatataaagtctctcccttggttggaagccaataatgctcttgacgtctc	749
D	570	GGAGAAAGAAATATATATAAGCTTTCCTGGTGGAAACCATACATATATGTCTGGCAGGTCT	629
Q	750	gaaagaagaatlgltccaaatggcctcttgacccactgtgcttlaactactaatacttcttca	809
D	630	GAAGAAGATGGTCCAAATGGGCTCTGACCACATGTGCTTACTACTACTATCTTTCTCA	689
Q	810	tgtgacctgtgtaaaagacttgaatcaggctcataltgagccctactagtatgtataga	869
D	690	TGTGACCTGGTATAAAGACTTTGAATTCAGGCCATCTTGAAGCCCTACTAGTATGTAGAGA	749
Q	870	aggaagctctggccaaggaagaagacacagcactgtgaccaatataactcttctgct	929
D	750	AGGAGATGTGGCCAAAGAAAGACACAGACTTTCGCAAAATTTATATCTTTTGTCTGT	809
Q	930	atttgaatgaaggagaagaatttggcacttaagaacaaagaactccttgaatgcagaatagga	989
D	810	ATTGTGATGAAGGGAAGAAGTTGGCACTCAGAAACAAAGAACTCTTGATGACAGATAGGA	869
Q	990	tgtcgtacatctgtctgggctctggccttaaatgcaacacatgcatatggttatgtataacagtc	1049
D	870	TGCTGCATCTCTCTGGGCTGGCTGTAAATGACACAGTCAATGGTATGTATTAACAGCTC	929
Q	1050	ctctgccaagctcttgatttgatgaccaagaagaataagctctatggcatgatatggaatggg	1109
D	930	TCTGCAAGTGTGATGTGATGCCACAGGAATATAGTCTATGGCATGTGATTTGGAATGG	989
Q	1110	caccaactcctgaatggaactcaatatacttcctcgaaagtcacaactctctgttgaagaaca	1169
D	990	CACCACTCTGTAATGCACTCAATATTTCCGGAAGTCAACACTTTCTGTGAGGAACA	1049
Q	1170	tcgccagcgctccttggaaatctgcacaactctccttactgctcaaacactcttgat	1229
D	1050	TGCGCAGCGCTCTTGGAAATCTCGCCAAATACTTCTTACTGCTCAACACTCTTGAT	1109
Q	1230	ggaccttggaaagtttactggttttgatatactcttccccaacaatagatgatgta	1289
D	1110	GGACCTTGGACAGTTTCTACTGTTTGTCTATCTTCTCCACCAACATGATGGCATGCA	1169
Q	1290	agctatgtcaaatcagaacgcgtctcagaaggagaccacaactcagaatgaaaaatcag	1349
D	1170	AGCTTATGTCAAATTAACACAGCTGTCCAGAGAACCCCACTACGAATGAAAAATATAGA	1229
Q	1350	agaagcggaaagactatgatagtacttactgatalcttgaatgatatgtgtlcataggttga	1409
D	1230	AGAAGCGGAAGACTATGATGATGATCTCTTACTGATCTGAAATGGAATGTGTCAGCTTGA	1289
Q	1410	tgaatgaacactctccttcccttataccaatctgctcagttgccaagaagatccccaac	1469
D	1290	TGATGACAACTCTCTCTTCTTATTCAAAATTCGCTCATGTTGGCAAGAACATCTTAAC	1349
Q	1470	ttgggtacatatactgtctgtcgaagaggagactgtactgtcccttatgctctgc	1529
D	1350	TTGGGTACATTACTTGTCTGCTGAAGAGAGACTGGGACTATGCTCTCCCTTAGTCTCTGC	1409
Q	1530	ccccatgacagaagttatataaagtcataattgaaacaaatggccctcagcagcgatgtgag	1589
D	1410	CCCCGATGACAGGAAGTTATTAAGAATCATATTTGAACAAATGGCCCTCAGCGGATGGTAG	1469
Q	1590	gaagtacaaaaaagctccgatttatgtgcatacacagatgaacctttaagactcgtlgaagc	1649
D	1470	GAACTACAAAAAAGTCGGATTTATGTGATACACAGATGGAACCTTTAAGACTCTGGAAGC	1529



OY	1650	tacacgacgaataacaggaatcttggacccttactcttaatggggaagtctggaacacact	1709
Db	1530	tatttcagatgaataacacaggaattcttggaccctttactcttaatggggaagtctggaacacact	1589
OY	1710	gtctattatattbaagaatacgaacagacacataataacatctacccctccaggaataac	1769
Db	1590	gttatattatattbaagaatacgaacagacacataataacatctacccctccaggaataac	1649
OY	1770	tgatgcgtccctcttgyatctcaagaggaattaccaaaaggtgttaaaaaacttgaagaattc	1829
Db	1650	tgatgcgtccgtcttgyatctcaagaggaattaccaaaaggtgttaaaaaacttgaagaattc	1709
OY	1830	tccaattctgcacggagaataatctcaatataaatggacatgactgtagaagaatggcc	1889
Db	1710	tccaattctgcacggagaataatctcaatataaatggacatgactgtagaagaatggcccc	1769
OY	1890	aacaaatcaagaatccctggcgtgcacccgcattacccttaagcttcgftaataatggaga	1949
Db	1770	aactaaatcaagaatccctggcgtgcacccgcattacccttaagcttcgftaataatggaga	1829
OY	1950	agactctagctctcaggaactcaatctggccctctccctctcactctgctcaaaagaatctgtaga	2009
Db	1830	agactctagctctcaggaactcaatctggccctctccctctcactctgctcaaaagaatctgtaga	1889
OY	2010	aagagggaacccagataatgtctcagaacaagagaatgtcalcctgttctctgtatttgatga	2069
Db	1890	aagagggaacccagataatgtctcagaacaagagaatgtcalcctgttctctgtatttgatga	1949
OY	2070	gaaccgaagcttggttacctcaacagaagaataaacaagcttctccccaatccagcttggaat	2129
Db	1950	gaaccgaagcttggttacctcaacagaagaataaacaagcttctccccaatccagcttggaat	2009
OY	2130	gcaagcttggaagatcccaagagttccaaagcctcccaacatctgacagacatcaaatggctatgt	2189
Db	2010	gcaccttggaagatcccaagagttccaaagcctcccaacatctgacagacatcaaatggctatgt	2069
OY	2190	ttttgataagtttgcagttgtcagtttqtttctgcaatgaagtgtgcatactgctcatctcaag	2249
Db	2070	tttttgataagtttgcagttgtcagtttqtttctgcaatgaagtgtgcatactgctcatctcaag	2129
OY	2250	cattggagacacagacggaactccctctctgtctctctctctgatactactcaaaacaa	2309
Db	2130	cattggagacacacagacggaactccctctctgtctctctctctgatactactcaaaacaa	2189
OY	2310	aatgctatagaaagacacactcaaccctatctcccatctcgaagagaacatgctctcaatgc	2369
Db	2190	aatgctctatagaaagacacactcaaccctatctcccatctcgaagagaacatgctctcaatgc	2249
OY	2370	gaatgaaccccaagctctcatatgaaattctggagtgtgcacaactcagaactcttggaacagag	2429
Db	2250	gatgaaccccaagctctcatatgaaattctggagtgtgcacaactcagaactcttggaacagag	2309
OY	2430	catgaacgctctcagaagtttctcagttgtgtgaagaagaactggttgatcttaacagga	2489
Db	2310	catgaacgctctcagaagtttctcagttgtgtgaagaagaactggttgatcttaacagga	2369
OY	2490	caagtaatgaagaatctctcagcactactgtctgtataaacaaatgcatcttgacaacaaga	2549
Db	2370	caagttatgaagaatctctcagcactactgtctgtataaacaaatgcatcttgacaacaaga	2429
OY	2550	ctctctcccaagaatctcaagacacccctagcaactgtgcacaaagaacatcttaatgcaccca	2607
Db	2430	ctctctcccaagaatctcaagacacccctagcaactgtgcacacaaagaacatcttaatgcaccca	2487

RESULT	13			
LOCUS	I31901	9009 bp	DNA	
DEFINITION	Sequence 3 from patent US 5583209			
ACCESSION	I31901			
VERSION	I31901.1			
KEYWORDS	GI:1822692			

SOURCE	Unknown.
ORGANISM	Unknown.
REFERENCE	Unclassified.
AUTHORS	1 (bases 1 to 9009)
TITLE	Lollar,J.S. and Runge,M.S.
JOURNAL	Hybrid human/porcine factor VIII
FEATURES	Patent: US 5583209-A 3 10-Dec-1996;
source	Location/Qualifiers
	1..9009
	/organism="unknown"
BASE COUNT	2853 a 1907 c 1844 g 2405 t
ORIGIN	
Query Match	49.0%: Score 2336.4; DB 10; Length 9009;
Best Local Similarity	100.0%: Pred. No. 0;
Matches 2337; Conservative	0; Mismatches 1; Indels 0; Gaps 0;
QY	270 catgcaaatagagctctccacccgtctctctctgctcttgagatctgcttaagtgc 329
DB	150 CATGCAATATAGAGCTCTCACCTGGCTTCTTCTGTCCTTTGGATCTGCTTTAGTGC 209
QY	330 caccagaagaatactaccctgggtgctgagtgagactgtcatgggaactatagcaaatgtatct 389
DB	210 CACCAAGAAGATACATCCCTGGGTGCAGTGGAACTGTCAATGGAGCATATATAGCAAAAGTATCT 269
QY	390 cggatagctgcctgtgagcgaagaattccctcctagatgtgccaatcttccatcaa 449
DB	270 CGGTAGCTGCTGTGGAGCGCAAGATTTCCTCTATGAGATGGCCMAAATCTTTCCATTTCAA 329
QY	450 caccctcagtcgtgtacaataaagaactctgttgttagaattcaogtltcaaccttcaacat 509
DB	330 CACCTCACTCGTGTACAAAAGACTCTGTTGTAGAAATTCACGGTTACCTTTTCAACAT 389
QY	510 cgcctaagccaagagccacccttgatggtgtctgctagtgctcctacatccagagctgaagttta 569
DB	390 CGCTAAGCCAAGGGCCACCTGGATGGGTCTCTAGTCTCTACATCCAGGCTGAGGTTTA 449
QY	570 tgateacagtgatcattacttaacttaagaanaatggcttcccatcctgtaagttctatgctgt 629
DB	450 TGATACACTGGTGCATTACACTTTAAGAACATGTGGCTTCCATCTGTCACTTTCATGCTGT 509
QY	630 tgggtatcctcacttgguaaagctctctgaaaggagctgaaatgatgatcatcagaccagtcaaa 689
DB	510 TGTGTATCTCTACCTGAGGAAGCTTCTGAGAGGAGCTGAATATGATGATCAGACACTCAAG 569
QY	690 ggaagaagaagaatgaataagctctccctggtgtggaagcatatcatatgcttgcaggtcct 749
DB	570 GGAGAAAGAAGATGATGATAAGTCTTCCCTGGTGAAGCCATATGATGCTGGCAGGTCT 629
QY	750 gaaagagaatggttccaatagtcctctgaaccacgtgtgctcttaacctatcatcttctctca 809
DB	630 GAAAGAGATAGTCTCAATAGGCTCTGACCCACTGTGCTTAACATCAATATCTTTCTCA 689
QY	810 tgttagacctgtgtaaaagagcttgaatctcaagggcctatcttgagccctactagatgtlaaga 869
DB	690 TGTGAGCTCTGGTATAAAGACTTGATTCAGGCCCTATTGAGACCCCTACTAGATGTAGAGA 749
QY	870 agggagctctgccaaggaanaagacacagacctgtgcacaaattatactactcttctgctgt 929
DB	750 AGGGAGTCTGGCCCAAGAAAGACACAGACTCTGCACAAATTTATCTACTCTTTTGGCTGT 809
QY	930 atttgatgaagggaanaagtttgacactcagaacaaagaactcccttgatgcaggaataga 989
DB	810 ATTTGATGAAGAAGGAAGATTGTGGCACTAGAAACAAAGAACTCCCTTGAATGCGATGAGTA 869
QY	990 tgcgtgcaatctctcggggccttgactaaatgacacagctcaatggttcatgtaaacaggtc 1049
DB	870 TGTGTGCACTCTCTGGGGCTTGCCCTAAATATGCACACAGCTCAATGTGTTATGTAAACAGGTC 929
QY	1050 tctgcacaggtctgatctgagtgccacaggaatacagtcctatgtgcatgtgatttgaatgg 1109
DB	930 TCTGTCCAGGTGTGATTTGGATSCCACAGAAATCAAGTCTATTTGGCATGTGATTTGAATGGG 989



QY	1110	caacactcctlgaagctgacatccaataatccctcogaaagtcacacattctctgtgaggaacca	1165
Db	990	CAACACTCCTCGAAGTGCACTCAATATTCTTCGAAAGGTACACATTTCTTGAGGAACCA	1049
QY	1170	tgcgaagagctctcttgaaatctcgcgaataactcttcttaactgctcgaacactctgat	1229
Db	1050	TGCGCAGGCGCTCTTGGAAATTCGCCAATTAACTTCTTACTCTCAACACTCTTGAT	1109
QY	1230	ggacacttggacaagttcttaactgttttgcatalctctccacccaataatga tggcatlga	1289
Db	1110	GGACCTTGGACACTTCTACTGTTTTGTCAATTCCTTCCACCACAATGATGAGCATGGA	1169
QY	1290	agctctatgcacaagctagacagcgttcgaagggaaccccaacttaagaatgaaataatga	1349
Db	1170	AGCTTATGTCAAAAGTGAAGACGCTGTCCAGAGGAACCCCAACTACGAATGAAAATAATGA	1229
QY	1350	agaagcggagaagactatgatcatcttaactgattcttcgaagtgaatgctgtgcacgtlga	1409
Db	1230	AGAGCGGGAAGACTATGATGATGATGATCTTACTGATTCGAAATGGAATGATGTGTCAGATTGA	1289
QY	1410	tgatgacaactctcctctcttcaatccaaltgcgtcagtttgcagaagaacatccaaac	1469
Db	1290	TGATGACAACACTCTCTTCTCTTATTCAAAATTCGCTCAGTTGCCAAGAAGCACTCTAAAC	1349
QY	1470	ttgggtacacttaactatgctgtctgaaggaggagcttggacatgctcccttaagctccgc	1529
Db	1350	TTGGGTACACTTACATTGCTCTGTAAGAGGAGAGATGGGACTATGCTCTTATGTCCTGC	1409
QY	1530	ccccga tgcagagaagtataaaatcaatattbaacaatggccctcagcgaattgtag	1589
Db	1410	CCCCGATGACAGAGATTATAAAAGTCAATTTGCAACAAATGCGCTCAGCGGATTTGGAG	1469
QY	1590	gaagtacaacaaaagctccgaattatlygcatacacagatlgaaaccttaagactgltgaagc	1649
Db	1470	GAAGTCAAAAAAGCTCGGATTTATGCGATCACAGATGAACCTTTAAAGCTCGTGAAGC	1529
QY	1650	tattcgcagctgaatcaggaatcttggacacttaacttta tggggaagt tggagacacact	1709
Db	1530	TATTTCAGCATGAATTCAGGAATCTTGGACCTTTATCTTATGCGGAAGCTTGAGACACT	1589
QY	1710	gttgaatataatlaagaatcaagcaagcagaacacataaacaatccactccacggaatcac	1769
Db	1590	GTTGATTATATTTAAGATCAACAGCAGACCATATTAACATCTAACCTCAGGAATCAC	1649
QY	1770	tgaatgcgctcctctgtatccaaggagatatcccaaaagtgtlaaacatttgaagaattc	1829
Db	1650	TGATGTCCTCTCTTATTATTCAAAGGAATTAACAAAAGGTAAACATTTGGAAGGATTT	1709
QY	1830	tccaattccgcagggagaaatattcaaaataaaatgaacagtgacgtlagaaga tggcc	1889
Db	1710	TCCAAATTCGCCAGAGGAATATTCAATATTTAAATGAGACAGTGAAGATGGGCC	1769
QY	1890	aactaaatcagaatcctcgtgtgcctgacccgctatlaactlagtttcgltlaatalgagag	1949
Db	1770	AACTTAATCAGATCCTCGGAGCTCGACCCCTATTACTTACTTACTTTCGTTAATATGAGAG	1829
QY	1950	agatctagcttcaagagctcatatggccctccctccatctgtctacaagaagaatctgtagatca	2009
Db	1830	AATCTAGCTTCAGGACTATTGCGCTTCCTCATCTGCTACAAAAGATCTTGATGATCA	1889
QY	2010	aagaggaacaeagataatgtcagacaagaagaatgtcatcctgttcttctlaatttga	2069
Db	1890	AAGAGGAACCAAGATATATGTCACACACAGAGGAATGTCTCTGTTTTCTGATTTGATGA	1949
QY	2070	gaaccgaagcttgtlactccaacagaaatalacaaagctttctccccaatccagcttga	2129
Db	1950	GAACCGAACCTGGTACTCTCACAGAAATTAACAACGCTTCTCCCAATCCAGCTGAGGT	2009
QY	2130	gaagcttggagatccagagtttccaagcttccaacatcatatgcaacagcatlaaatlygctatgt	2189
Db	2010	GAAGCTTGGAGATCCAGAGTTCCAAAGCTTCCAAATCATGATCAGACAGCATCAATGGCTATGT	2069

QY	2190	ttttatagattcgcaagttgtaagtttgtttgtaatgaagtgagcaatccggttaattctaa	2249
Db	2070	tttttatagatttgacagttgttcagattgtttgatatgagtgagcattcggtatattcttaag	2129
QY	2250	cattgagacagactgactcttccttcgtctctctctctctgtataactcttaacacaa	2309
Db	2130	catttgagacagactgactcttccttcgtctctctctctgtatattaccttcaaacacaa	2189
QY	2310	aattgtctatgaagacacactcaacctatcccatctctcagaagaagactgtcttcatg	2369
Db	2190	aattgctatgaagacacactcaacctatcccatctctcagaagaagactgtcttcatg	2249
QY	2370	gattgataaaacccaggtctctatgtaattctgggtgtccacaacactaagacttcggaagaag	2429
Db	2250	gattgataaaacccaggtctctatgtaattctgggtgtccacaacactaagacttcggaagaag	2309
QY	2430	cattgacccgctctacgaaggtctctatggtgtgtacagaagacactgtgtattatcagaga	2489
Db	2310	cattgacccgctctacgaaggtctctatggtgtgtacagaagacactgtgtattatcagaga	2369
QY	2490	caattatgaagaattctcagcaactgtgtcgtgaataaacaatgtccattgaaccaagaag	2549
Db	2370	cagttatgaagaattctcagcaactgtgtcgtgaataaacaatgtccattgaaccaagaag	2429
QY	2550	ctctctccagaattcaagaacacccttagcaactagcagaagaacttaattgaaccacca	2607
Db	2430	ctttctccagaattcaagaacacccttagcaactagcagaagaacttaattgaaccacca	2487
RESULT	14		
LOCUS	163424	163424	9009 bp DNA
DEFINITION	Sequence 1 from patent US 5663060.		
ACCESSION	163424		
VERSION	163424.1	GI:2480997	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 9009)		
AUTHORS	Lollar,J.S. and Runge,M.S.		
TITLE	Hybrid human/animal factor VIII		
JOURNAL	Patent: US 5663060-A 1 02-SEP-1997;		
FEATURES	Location/Qualifiers		
source	1..9009		
BASE COUNT	2853 a	1907 c	1844 g 2405 t
ORIGIN			
Query Match	49.0%; Score 2336.4; DB 10; Length 9009;		
Best Local Similarity	100.0%; Pred. No. 0;		
Matches 2337; Conservative	0; Mismatches	1; Indels	0; Gaps
QY	2710	catgcaaatagagctctccacactgtctctctctgtgaccttttgcgaatttcgatactgtaagtgc	329
Db	150	catgcaaatagagctctccacactgtctctctctgtgaccttttgcgaatttcgatactgtaagtgc	209
QY	330	caaccgaagatactactcctgggtgtcagtggaatgtcattggagactatatacgaaagtgaact	389
Db	210	caaccgaagatactactcctgggtgtcagtggaatgtcattggagactatatacgaaagtgaact	269
QY	390	cggtagagctgcacctgtggagcgcgaagaattctccctcctagtagtgccaaaattcttccattaa	449
Db	270	cggtagagctgcacctgtggagcgcgaagaattctccctcctagtagtgccaaaattcttccattaa	329
QY	450	caaccctagctcgtgtacaaaagaactctgttttgtaaatctacoggtgtcacaccttlttaacat	509
Db	330	caaccctagctcgtgtgtacaaaagaactctgttttgtaaatctacoggtgtcacaccttlttaacat	389
QY	510	cgttaagccaagagcaccctgtgattggtctgtctaggtccataccalccaaagcttgagttta	569
Db	390	cgttaagccaagagcaccctgtgattggtctgtctaggtccataccalccaaagcttgagttta	449



QY	570	tgaatacagtggtcattacacttaagaacatggtctcccaatccctgtccagctctcatgctgt	629
Db	450	TGATATACAGTGGTCAATTACCTTAAGAAACATGGCTTCCACTGTCAGTCTTCATGCTGT	509
QY	630	tgtgtatcctacttgaaagctctctgagggagcttgaaataatgataatgcagacagctcaag	689
Db	510	TGtGTATCTCTACTGGAAGACCTTCGAGGGAGCTGTAATGTATGTATCAGACCACTCAAG	569
QY	690	ggaaagaagaatgataaagaatctctccctgtgtgaaagccaataatatgctctgtgagctct	749
Db	570	GGAAGAAAGAAAGATGATTAAGTCTTCCCTGGTGGAAACCCATTAATATGTCTGGACAGTCT	629
QY	750	gaaagagaatggtccaaatggtccctctgacccacagtgtcccttaactaactaatctctcca	809
Db	630	GAAGAGAAATGATGTCATGAGCCCTGTACACCACAGTGTCCCTTACTACTCATATCTTTCGA	689
QY	810	tgtgagccctgtlaaagaacttgaaatcgaagccctcatltgagccctactagtaatgaaaga	869
Db	690	TGtTGACCTGTGTAAGAACTTGAAATTCAGGCCCTCATTTGAGGCCCTACTAGTATGAGAGA	749
QY	870	agggagctctgtccaaggaagaagacaagagcccttgacaaattataactacttltgtgctgt	929
Db	750	AGGGAGTCTGTCCGCAAGGAAAAAGACACAGACTTTGCAAAATTTAATCACTTTTGTGCTGT	809
QY	930	attgtagaagggaagaagtgttgacccagaagaacaagaactccttgatgagcagataagga	989
Db	810	ATTTGATGAAAGGAAAAAGTTGGCAGCTCAACAAACAAAGAACTCTTGATGTCAGGATAGGA	869
QY	990	tgtctgcatactgtctcgtggtccgtgacctlaaaaatgacaagaatgaaatgtgtlaatglaaagaagtc	1049
Db	870	TGCTGCATCTGTCTCGGGCCGTGGCTTAATATGCACACAGTCATGGTATTATGAACAGTCT	929
QY	1050	tctgtccaggtctctgatttgatgatagcacaagaagaatcaagtctatctgtgcaatgfatgtgaaatgg	1109
Db	930	TCTGTCCAGGCTCTATTGTGAAGTCACACAGGAATACAGTCTATTGGCATGTGATGGATGG	989
QY	1110	cacacacccttgaaagtgaacccaatatctctctgaaggtacacattctctgtgagaagaaca	1169
Db	990	CACCACTCTCGAAGTGCACCTCAATATTCTCGAAGTCCACACATTTCTTGAGAGAACCA	1049
QY	1170	tcgcagagcgctctcttggaaaatctcgccaataactctccctactctgctcaacaactctgat	1229
Db	1050	TGCGCAGCGCTCCCTGGAATCTCGCCAAATCTTCCCTACTGCTCAACACTCTTGAT	1109
QY	1230	ggaaccttgagacagttctctactgttctgtgcatatctctcccaacaatgataatgcatgaa	1289
Db	1110	GGACCTTGGAAGATTTCTACTGTTTGTATATCTCTTCCACCAAACTGATGGCATGGA	1169
QY	1290	agcctatgataaagtgaagcagctgtctcagaagaagcccaacacagaatgaaaaataatga	1349
Db	1170	AGCTTATGTCAAAAGTGAACAGCTGTCACAGGAAACCCCACTACGAAATGAATAAATGA	1229
QY	1350	agaaagcggaagactatgatgatgatctctactctgatactctgaaatgaaatgtgtgtcgaagtgttga	1409
Db	1230	AGAAAGCGGAAGACTATGATGATGATGATCTTACTGATCTTGAAATGATGTGTGACAGTTGA	1289
QY	1410	tgatgatcaactctcccttcccttataccaatctgcgtcagltgtccaaagaagactcctaaac	1469
Db	1290	TGATGACAACTCTCTCTCTTTATCTCAAAATTTGGCTCAGTTGCCAAGAGACATCTCTAAAC	1349
QY	1470	tttggatcatataatgtctgtctgaaagaagaagacgtggactatgctcctctttagctctgc	1529
Db	1350	TTGGGTACATTTAATTTGCTGTCTGAAGAGGACTGTGGGCTATGTCTCTCTTATGCTCTGC	1409
QY	1530	ccccgatgacagaagtatataaagatcaaatatgtgacaaatggtccctcagcgatgtgtgag	1589
Db	1410	CCCCGATGACAGAAAGTTATTAAGTCATAATTTTGAACAATAGCCCTCAGCGGATTTGTAG	1469
QY	1590	gaagatacaaaaagtcgcgatttatgtgcatacacagatgaaacctttaagactcgtlgaagc	1649
Db	1470	GAAATACAAAAAAGTCGGATTTGTGGCATATACAGATGAAACCTTTAAGACTCTGGAAAGC	1529

[illegible]

RESULT	15
HOMFVIII	
LOCUS	HOMFVIIII 9009 bp mRNA PRI 08-NOV-1994
DEFINITION	Human coagulation factor VIII:C (anti-hemophilic factor) mRNA.
ACCESSION	K01740
VERSION	K01740.1 GI:182802
SOURCE	factor VIII; hemophilic.
REMARKS	Human cDNA to liver mRNA, and genomic DNA.



ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
REFERENCE 1 (bases 1 to 9009)  
AUTHORS Toole,J.J., Knopf,J.L., Mooney,J.M., Sultzman,L.A., Buecker,J.L.,  
Pittman,D.D., Kaufman,R.J., Brown,E., Shoemaker,C., Orr,E.C.,  
Amphlett,G.W., Foster,W.B., Coe,M.L., Knutson,G.J., Fass,D.N. and  
Hewick,R.M.  
TITLE Molecular cloning of a cDNA encoding human antihemophilic factor  
JOURNAL Nature 312 (5992), 342-347 (1984)  
MEDLINE 85061550  
COMMENT Currently hemophiliacs are treated with factor VIII:C purified from  
blood. The human factor VIII:C cDNA has now been isolated, cloned  
and expressed. The gene was also isolated. It consists of a large  
number of exons spanning over 180 kb. The authors kindly supplied  
the sequence directly to Genbank by computer network.  
FEATURES  
source location/Qualifiers  
1..9009  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/map="xq28"  
151..7206  
/gene="F8C"  
151..7206  
/gene="F8C"  
/note="factor VIII"  
/codon\_start=1  
/db\_xref="GDB:G00-119-124"  
/protein\_id="AA52484.1"  
/db\_xref="GI:182803"  
/translation="MTELSLTCFLCLLRFCFSATRRYVGLAVELSDYMSDILGELP  
VDARPPRVKSEPTNVVYKKTLEFVYHLINLAKPRPMWGLGPTLOAVYDVT  
VITLKNMASHPVSLAHGVSYMKASGAEVDDSDTSREKEDVVFGEHTYVWYL  
KENGPMASDPLCLLYSYLSHVDLYKDLNSGLIGALVCRESLAKEKQTQTHKFLLE  
AVPDECKSMSETRKNSMODRNASARAMPKRVNKNVNSPLSLGICGCHKSYVWH  
IGMCTPSEHSIELECHTPIVNRNROASLESPITPTAQLMLDLGLOLPCHTSSH  
QHDGMEAYAVDSCPEEPQLRMKNNEAEEDDDLTSEMDVVFDDNSFSIQIRS  
VAKHPKTVWYIAAEEDMDYALVLAPODRYSOYLNNGPQIRGRKRVFMAY  
TDETFKTRALIQHSGILGLVGEVDTLIRKNQASREYNINYPGIDIVRLYSR  
RLPGVYHLKDFPLPGEIFRKWTYVEDQPTKSDPCLRLRYTSEFNMDASGL  
IGPLLCYKESVDORQNOIMSDKRNVLTFSEFDRNSWYLENIOFLPNAAGOLE  
PEROASIMHSINGVPSLOSYCLHEVYMYLISGAQDFLTSVSPSCPRHKVY  
YEDTLTPFPGSETVPMSENGMLITGCHNSDRNRMQLLWVSSCDKGTGTED  
SYEDISAYLLSKNNALEPRSFQNSRHPSTRQKOFNATTPENDIEKTDWFAHTPM  
PKIQNVSSDLMMLROSPTPHGLSLDLQAEKETEESDPSGAIDSNNSLSMTHE  
RPLHSGDWVFTPEESGLQRLNEKLTGTATTELKLDKFKVSTSNLITIPSDNLA  
AGTDNNTSLGPPSPVHYDSOLDPTLFGKSSPLTEGSGPLSEENNDSSLSGLM  
NSQSSMGKVVSTESGRLFKRAHGAPLITKQNALPKVSTILKTKNKSNNATNR  
KTHIDPRLIENSPSYMONILBESDTEFKVTPLIHRMLDKNATRLRLNHSNKTT  
SSKNMEVQOKKEGPIPPDAONPDMSFFKMLTFESARWLTQRTGKNSLNGQSPSK  
QVLSGPEKSEVGEONLSEKNVYVVGGEFTDGLKEMVPPSRNLFLNLMJHEN  
NTHNOEKKIQEEIEKETTLEQENVLPQITVTGKNPKMKLFLSTRONWEGSYEGA  
YAPVLODFRLINDSTNRTKHTAHFSKGEENLEGNQTKOIVEKYACTTRISPT  
SOONPVTORSKRALOKOPLLETELEKRLITVDTSQWMSNMKHLPSITLODYNM  
KEGALITQSPSLSDCLTRSHSIPQANRSPPLAKVSSPITPILTRYLRONDSHP  
AASYRRKDSGVQSSHPFLGQAKNNLSLALITLLETWDQREVSGLSATNSVYKYV  
ENTVLPKRPDLPTKSGVELLPKVHLYQKDLPTPTSGSPBHLVGLSLQGEKLA  
KNMEANRPGKVPFLRVATBESSAKTSPKILPLADNHYGQIPEEKMSQEKSEKTA  
FKKQDTLSLNACESNHAIAINEGONKPELEVYMAWAGRGRETSQNPPLAKRELM  
ITRTLOSODEEIDYDITISVEMKKEDEDDIYDEDNQSPREFOKKTTHFLTAVERLM  
DYGMSSSPHYLRNRAQSGVPOKKVYFOETDQSFOPILRGLENLGLGTYIRA  
EVEDNITWYTRNQASRPSYSSLSISTEEOQGAERKKNVKNETKYFWQOHHA  
APTKDEFCAMAYFSDVDLEKDVHSGILPLACHNTLNPAGROVYTOEFLPFT  
IPEDETSWYFTTENNERCARPCNQIOMEDPEFKENYRHALINGYIMDTLPGLVMAQOR  
IRMYLTSMGSENENHSHFSGHYETVVRKKEEKYKALNLVPGVEYEMLSKSGIMR  
VECLIGEHLAGMSTLELYVSNKQPTPLGASGIIIPQITASQYQOMAPKLRLHY  
SGSINAMSTREPFSWIKVDLAPNITIGITIGQARQFSFYSISQITLMSLDKMKQ  
TYRGNSTGTLMEFEGVNDSSGIRKHNINPPIARYIRLHPHYSIRSTLRNMLGCDL  
NSCSMPDGMESKASIDAOITRASSYFTNMFATWSKARLHLOGSNMARPVNNPKEM  
LAYDFQTKMKGAVTGTQGVKSLTSMYKFEFLSSODGHWMTLFPNGKKVYQNGNO  
DSFTPVNSLDPLTRYLRIRIHQPSWHLALRNEVLGCEAODLY"

BASE COUNT 2853 a 1907 c 1844 g 2405 t  
ORIGIN 160 bp 5' to 3' cut site; chromosome Xq28.

Query Match	49.0%; Score 2336.4; DB 97; Length 9009;
Best Local Similarity	100.0%; Pred. No. 0;
Matches 2337; Conservative	0; Mismatches 1; Indels 0; Gaps 0;
Oy	270 catgcaaatagagctctccaccctgctctctctctgctcttgagcttgcttagtgc 329
Db	150 CATGCAAAATAGAGCTCTCCACCCTGCTCTCTCTGCTTGGCAATCTCTTAACTGTC 209
Oy	330 caccgaagatactaccctgggtgacgtggaactatcatgggaactatgcaaatgactc 389
Db	210 CACCGAAGATACATACCTGGGAGCATGGAACCTGATCAGGAGCATATAGCAAACTGATCT 269
Oy	390 cgttagagctgcctggagcgaagattccctccctagaagtgccaactcttccatcaa 449
Db	270 CGGTAGAGCTGCTGTGACGACGAAGATTTCCTCTAGAGATGCCAAATCTTTCCATTCAA 329
Oy	450 caacctcagtcgtgtacaaaagaactctgtctttagaattcaacggtlccacclttcaacat 509
Db	330 CACCTCAGTCTGTCATCAAAAAGACTCTGTTGTAATTCACGGTTCACCTTTCAACAT 389
Oy	510 cgtcgaagcgaagggccacccttgatgagtcctgctagtctcctaccatccaagctgaggttta 569
Db	390 CGCTAAGCCAAAGGCCACCTTGATGGTCTCTAGTCTCAACATCCAGGCTGAGGTTTA 449
Oy	570 tgatacagtgatcatctacacttaagaacatggtctcccatcctgtaagctctcagtcgtct 629
Db	450 TGATCAGTGGTCATTTACACTTAAAGACATAGCTGGCTCCATCTCTCATGTCTGT 509
Oy	630 tgggtatccctacttgaaagctctcgaaggagctgaaatagatgacagaccagtcagaag 689
Db	510 TGGTATACCTACTGTAAGAGCTTCTGAGGAGGAGCAATATATATATATAGACCACTCAAG 569
Oy	690 ggaagaagaatgataagatgctctccctgggtggaagccatcatatgctctggagagctct 749
Db	570 GGAGAAAGATGATTAAGTCTCTCCCTGGTGAAGCCATACATATGCTGGCAGGTCCT 629
Oy	750 gaaagaagaatgctcccaatgagctctgacccactgtgacctaccatcatatctcttca 809
Db	630 GAAAGAAGATGATGCAATAGGCTCTGACCCACTGTGCTTACATATATCTTCTCA 689
Oy	810 tgtggaactggtataaagacttgaattcaaggctcatatggagccctactagatagtaaga 869
Db	690 TGTGACCTGTGTAAGAACTGGAATTCAGGCTCATATGAGCCCTACTAGATGTAAGA 749
Oy	870 agggagcttgcccaagaaagaacaagaactctgcaaaatttatctacttcttgctgt 929
Db	750 AGGGAGCTGGCCAGAAAGACACAGACCTTGACAAATTTATACTATTCTTTGCTGT 809
Oy	930 atttataagaaggaaagtctggcactcagaaacaaagaactccttgatgacagataagga 989
Db	810 ATTTGATGAAGGAAAGAACTTGGCACTGAAACAAAGAATCTTGAAGCAGATAGGA 869
Oy	990 tgcgtgacatcgtctgggcttgagcttaaaatgacacagtgtaatgtaaacaggtc 1049
Db	870 TGCTGCATCTGCTGGGCTTGCCCTAAATATCACACAGTCATGTTATGTAACAGATC 929
Oy	1050 tctgccaagcttgatgtagatgccacaggaataagctcatatggagatgtagtggaaatgg 1109
Db	930 TCTGCCAAGCTTGATTTGATGATGCCACAGGAATATCATGCTATTGGCATGTGATTTGAATGGG 989
Oy	1110 caccactctgaagtgcactcaatatctctcgaagagtcacacattctcttgaggaacca 1169
Db	990 CACCACTCTGTAAGTGACCTCAATATCTCTGGAAGTGTCACACATTTCTTGGAGAACCA 1049
Oy	1170 tcggcagagcgtctccttggaatctcgcgaataaacttcccttaactgctcgaacactctgat 1229
Db	1050 TCGCCAGGCGTCCTTGGAAATCTGCCAATACCTTCTTACTCTCAACACACTCTTGAT 1109
Oy	1230 ggaacttggaagatttcaactggttttgcatatctctccacaacaatgtaggacgga 1289
Db	1110 GGACCTTGACAGATTTTCTACTGTTTGTCAATATCTTCCACCAACATGATGGCATGGA 1169







**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 18, 2001, 04:50:04 ; Search time 389.82 Seconds  
(without alignments)  
2566.193 Million cell updates/sec

Title: US-09-689-430-1\_COPY\_419\_4835

Perfect score: 4417  
1 catgcaatagagctctcca.....aggcacagcactctactga 4417

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapect 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_NA.\*

1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS.COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/Backfile1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysts of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
✓ 1	4412.8	99.9	4629	2	US-08-484-891-7 Sequence 7, Appli
✓ 2	4407.4	99.8	9354	1	US-08-683-839B-2 Sequence 2, Appli
✓ 3	4357.4	98.7	4670	3	US-08-717-294-41 Sequence 41, Appli
✓ 4	4319.8	97.8	4999	4	US-09-470-618-14 Sequence 14, Appli
✓ 5	4319.8	97.8	4999	4	US-09-364-862-14 Sequence 14, Appli
6	4249.4	96.2	11933	4	US-09-470-618-13 Sequence 13, Appli
7	4249.4	96.2	11933	4	US-09-364-862-13 Sequence 13, Appli
8	3743.8	84.8	5035	2	US-08-882-083-1 Sequence 1, Appli
9	3743.8	84.8	5035	2	US-08-558-107-1 Sequence 1, Appli
10	3743.8	84.8	5035	2	US-09-243-539-1 Sequence 38, Appli
11	3189.8	72.2	4334	2	US-08-670-707A-38 Sequence 38, Appli
12	3189.8	72.2	4334	4	US-09-037-601-38 Sequence 38, Appli
13	2603.8	58.9	4451	3	US-08-717-294-42 Sequence 42, Appli
14	2336.4	52.9	9009	1	US-07-864-004B-3 Sequence 3, Appli
15	2336.4	52.9	9009	1	US-08-212-133A-3 Sequence 3, Appli
16	2336.4	52.9	9009	1	US-08-212-133A-1 Sequence 1, Appli
17	2336.4	52.9	9009	1	US-08-474-503-1 Sequence 1, Appli
18	2336.4	52.9	9009	2	US-08-670-707A-1 Sequence 1, Appli
19	2336.4	52.9	9009	4	US-09-037-601-1 Sequence 1, Appli
20	2336.4	52.9	9009	5	PCT-US93-03275-3 Sequence 3, Appli
21	2336.4	52.9	9009	5	PCT-US94-13200-1 Sequence 1, Appli
22	2336.4	52.9	7056	1	US-08-121-202-1 Sequence 1, Appli
23	2333.2	52.9	8241	6	5171844-1 Patent No. 5171844
24	2333.2	52.9	8967	1	US-08-366-851A-1 Sequence 1, Appli
25	2275.2	51.5	6999	1	US-08-276-584A-1 Sequence 1, Appli
26	1848	41.8	7032	4	US-09-324-867-1 Sequence 1, Appli
27	1801.8	40.8	7493	1	US-08-212-133A-7 Sequence 7, Appli

28	1801.8	40.8	7493	1	US-08-474-503-5 Sequence 5, Appli
29	1801.8	40.8	7493	2	US-08-670-707A-5 Sequence 5, Appli
30	1801.8	40.8	7493	4	US-09-037-601-5 Sequence 5, Appli
31	1801.8	40.8	7493	5	PCT-US94-13200-5 Sequence 5, Appli
32	1751.2	39.6	6402	2	US-08-670-707A-36 Sequence 36, Appli
33	1751.2	39.6	6402	4	US-09-037-601-36 Sequence 36, Appli
34	1240.4	28.1	1623	1	US-08-121-202-3 Sequence 1, Appli
35	893.4	20.2	1130	1	US-07-864-004B-1 Sequence 1, Appli
36	893.4	20.2	1130	1	US-08-251-937A-1 Sequence 1, Appli
37	893.4	20.2	1130	1	US-08-212-133A-5 Sequence 5, Appli
38	893.4	20.2	1130	1	US-08-474-503-3 Sequence 3, Appli
39	893.4	20.2	1130	2	US-08-670-707A-3 Sequence 3, Appli
40	893.4	20.2	1130	4	US-09-037-601-3 Sequence 3, Appli
41	893.4	20.2	1130	5	PCT-US93-03275-1 Sequence 1, Appli
42	893.4	20.2	1130	5	PCT-US94-13200-3 Sequence 3, Appli
43	301.8	6.8	6909	2	US-08-804-186-1 Sequence 1, Appli
44	301.8	6.8	6909	2	US-08-658-340-1 Sequence 1, Appli
45	301.8	6.8	6909	3	US-08-746-111-26 Sequence 26, Appli

## ALIGNMENTS

RESULT 1  
US-08-484-891-7  
Sequence 7, Application US/08484891  
Patent No. 5935935  
GENERAL INFORMATION:  
APPLICANT: Connelly, Sheila  
APPLICANT: Kaleko, Michael  
APPLICANT: Smith, Theodore  
TITLE OF INVENTION: Adenoviral Vectors for  
TITLE OF INVENTION: Treatment of Hemophilia  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Carella, Byrne, Bain, Gilfillan,  
ADDRESSEE: Cecchi, Stewart & Olstein  
STREET: 6 Becker Farm Road  
CITY: Roseland  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07068  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch diskette  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WordPerfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/484,891  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/218,335  
FILING DATE: 25-MAR-1994  
APPLICATION NUMBER: 08/074,920  
FILING DATE: 10-JUN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Olstein, Elliot M.  
REGISTRATION NUMBER: 24,025  
REFERENCE/DOCKET NUMBER: 271010-273  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-994-1700  
TELEFAX: 201-994-1744  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4629 bases  
TYPE: nucleic acid  
STRANDEDNESS: single  
MOLECULE TYPE: linear  
FEATURE:  
NAME/KEY: Factor VIII CDNA with  
NAME/KEY: B domain deleted







OY	2102	atgtaaaacccaggctctatgtaattctggggtgcacaaactcaagcttcctcgaaacagagc	2101
Db	2101	ATTGAAAACCCAGGCTCTATGTGATTTCTGGGGGTGGCCAACTCGAGCTTTGGGAACGAGAGC	2160
OY	2162	atgacgccttactgaaggtttctagtttttgcaagaacacgtggtattatcaagagc	2221
Db	2161	ATGACCGCCTTACTGAAGGTTTTCTAGTTGTGACAGAAACACTGGTGATTTATTCAGAGAC	2220
OY	2222	agttatgaagatatltcoagatactgtcgtgagtaaaacaatgcatgtcaatgaaacaaagac	2281
Db	2221	AGTTATGAATATTTTCAGCATCTGTGCGAGTAAATAAATAGCCATTGAACCAAGAAGC	2280
OY	2282	ttctccagaatltcoaagacccctagcactagycaaaagcaattaatgcaaccaccaca	2341
Db	2281	TTCTCCAGAAATTTCAAGACACCCCTAGCACTAGGCAAAAGCAATTTAATGCCACCCACCA	2340
OY	2342	gtcttgaagaagcatataaagggaaataactgttactactcttcagttcaatgtaagaaggaa	2401
Db	2341	GTCTTTGAAGCCCATACACGGAAATACOTCTACTACTCTTCACTCGATGATCAAGAGAA	2400
OY	2402	attgacatgatbataccatcatcagtttgaatgaaagaagaagatlttgacattatgat	2461
Db	2401	ATTGACTATATGATATACCATATATAGTTGAATGAAGAGAGAGATTTTGACATTTATGAT	2460
OY	2462	gaggatgaanaatcagagcccccagctcttcaaaagaanaacagcacatttlatgtc	2521
Db	2461	GAGGATGAANAATCAGAGCCCCCGCACCTTTCAAAAAGAAAACGACACTATTTTATGCT	2520
OY	2522	gcagatgagaagctcttgggaattatggaatgagtagctcccaacatgctcttaagaacag	2581
Db	2521	GCACTGGAGAGGCTCTGGGATTTAGGGAAGAGTAGCTCCCACTCAATGTTCTTAAGAAACAG	2580
OY	2582	gtctaaagtgagagtttccctcaatgttcaagaagaattgttttccaagaattttacgatgagc	2641
Db	2581	GCTTAGAGTGGCAAGTGTCCCTCACTTCAGAAAGATGTTTTCAGAGAAATTTACTGATGCG	2640
OY	2642	tccttctacagcccttataccgttggagaactcaaatgaaacatttggagctcctggggcca	2701
Db	2641	TCCCTTACTCAGCCCTTATATCGGTGAGAACTTAATGAACAATTTGGACTCCTGGGGCCA	2700
OY	2702	tataaagacagaagttgaaagtaataatcatgtaactttcaagaatcaggcctctcgt	2761
Db	2701	TATATTAAGACACAGAAGTTGAAGATTAATATCATGTAACCTTTGAGAAATCAGGCCCTCGT	2760
OY	2762	cccatctccttcatctotagcccttatttttttgggaagatccaagggcaagggagcagaa	2821
Db	2761	CCCATTTCTTTATTTCTTACCTTATTTTCTTATGAGAGGATCAAGGCAAGAGACAGAA	2820
OY	2822	cctgaaaaaaacttctgaagccttaatgaaaccaaacttacttcttggaaagtccaacat	2881
Db	2821	CTTGAGAAAAAATTGTGTCAAGCCTAATGAACCAAAACTACTTTTGGAAAGTGCACAT	2880
OY	2882	cataatgaccccaataaagaatgattgttagctgcaaaagccttggccttattctctgaagt	2941
Db	2881	CATATGCGACCCCACTAAAGATGAGTTTGACTGCAGAAAGCCTGGGCTTATTTCTGATGTT	2940
OY	2942	gaacttgaaaaaagatgtgcaactcagcgcctgatgtgaacccctctggtctgccaactaac	3001
Db	2941	GACCTGGAAAAAATGTGTGCACTTAGGCCCTGATTTGGACCCCTTCTGTGTGGCACACTAAC	3000
OY	3002	acacttgaaccctgctatgtggagaacaagtgaagcaagtaagaatltgtgccttlttcaac	3061
Db	3001	ACACTGAACCTCTGCTCATGTGGAGACAAAGTGAACATGACAGAAATTTGCTCTGTTTTTCAC	3060
OY	3062	atctcttgaagaccaaagctgtgtaacttcaacttgaataatataatgaaagaacttgcagggct	3121
Db	3061	ATCTTTGATGAGACAAACAGCTGTGTACTTCACTGAAGAAATATGGAAGAAGAACGACGGCT	3120
OY	3122	ccctgtcaatatccagatgtaagaatcccacttttaagagaattatcgcttccatgtaac	3181
Db	3121	CCCTGCAATTTCCAGATGAGTAAGATCCCACTTTTAAAGAAATATATGCGTTCATGCAATC	3180

OY	3182	aaigycfaacataatggaataacactgctctgaatgaatgctccgaatcaagaattcga	3241
Db	3181	AATGGCTACATAAATGAATACACTGCTGCTTAAGTAATGGCCAGAGATCAAGGATTCGA	3240
OY	3242	tgatcatctgcacgaatgggcgaacgaatgaacaacatcatctcatcttaacgtggacat	3301
Db	3241	TGGTATCTGCTCAGCATGGGCAGCAATGAAMACATCCATTCTATTTCATTCAATGAGACAT	3300
OY	3302	gtgtcactcgttacgnaaaagaagagatlaaaatgycactgtacaatctatccagtt	3361
Db	3301	GTGTTCACCTGACGAAAAAAGAGAGATTAATAATGGCACTGTACAAATCTGATCCAGCT	3360
OY	3362	gtttttgaagacgttgaaatggttacatccaaagctgtgaatttggcgggttgaaatgctt	3421
Db	3361	GTTTTTSAGACAGTGGGAAAGTTTACCATCCAAAGCTGGAATTTGGCGGGTGGAAATGGCTT	3420
OY	3422	attgggaagacatacaagttgttggaatgaacactttttctgtgttgaaagaaatgaatgt	3481
Db	3421	ATTGGGACGACTTACACTGTGTGGAAATGACACTTTTCTGTGTGATACGCAATTAAGTGT	3480
OY	3482	cagagcccccgtggaatgagctctctggacacatlaagatattcgaattacagctctcaaga	3541
Db	3481	CAGACTCCCTCGSSAATAGGCTTCTGGACACATTAGAGATTTTCACATTACAGCTTCAGGA	3540
OY	3542	caataatgaacagttggccccaagaactgycocgaatcatalatccggatcaatcaatgccc	3601
Db	3541	CAATATGAGACAGTGGGCCCAAGAGCTGGCCAGACTTCATTATTCGGATTCATCAATGSCC	3600
OY	3602	ttggagcaaccaagagagcccttttcttgatcaaggttgatctgtttggccaatgatattt	3661
Db	3601	TGGAGCACCAAGGAGCCCTTTTCTTGATTCACAGTGTGATCTTTTGCCACCAATGATATTTT	3660
OY	3662	caaggaatcaagaaccagagggtgcacgtctcaagaattctccagctctcaactctcgaatt	3721
Db	3661	CACGGCATCAAGACCCAGGGTGGCCGTGCAAGATTTCTCAGGCTCTACTCTCTAGTTT	3720
OY	3722	atcaatcgtatagttcttgaatggnaaagttggcagaattatcgsgaataatccactgga	3781
Db	3721	ATCATCATGTTATGATCTTTGATGGGAAGATGGCAGCACTTATCGAGGAATTCACCTGGA	3780
OY	3782	acctaatggtctctcttggcaatlgtygatttaactcgggataaaacaatatatttaac	3841
Db	3781	ACCTTAATGCTCTTCTTTGGCAATGTGGATCTACTCGGGATTAACACAAATATTTTTTAAC	3840
OY	3842	ccccaatattgtctgcatacatccgttttggaccaacacatcatatagatctcgagcaact	3901
Db	3841	CCTCCAATTATGTGTCGATATCATTCGCTTTGGCAACCACATCATTAATGCAATTCGACGACT	3900
OY	3902	cttcgcacatgagatgtaatggtctgttgatattaaatgattgcaagatgccatlggnaatgag	3961
Db	3901	CTTCGCAATGGAGTTGATGGGCTGTGATTAATAATAGTTGCAGCATGCCATTTGGGAATGGAG	3960
OY	3962	agtaagaacataatcagaatgcacagatatctgtctcaactctcaacttaacaatatgtttgcc	4021
Db	3961	AGTAAGCAATATCATAGATGACAGATTAATCTGCTTCACTTATTCACAAATATGTTTTGCC	4020
OY	4022	acctggtctctcttlaaaaagcttgactcaactccccaaggagagatgaatgacctggagacct	4081
Db	4021	ACCTGGTCTCTCTTAAAGCTCCACTTCACCTTCCAAGGAGGAGATTAATGCTCGAGACTT	4080
OY	4082	caggatgaataatccaaaagaagtgtctgtcgaatggaatcttcacagaagacaaatgaagtcca	4141
Db	4081	CAGGGAATATATCCAAAAGATGGCTGCAAGTGGACATTCAGAGACAAATGAATGTACA	4140
OY	4142	ggagtaactactcaggagatgaaaatctctgtctacacgacatgtatgtgaaggagttcttc	4201
Db	4141	GGAGTAACACTACTCAGGAGATGAATATCTCTGCTTACAGCATGATTAATGTAAGAGATTCTCTC	4200
OY	4202	atcccaagagatcaaaatggtgcataagtggaactcttttttcagaaatggaagaatgaag	4261
Db	4201	ATCTCCAGCAGCTCAAGATGGCCATCATGATGAGACTCTCTTTTTCACAAATGAGCCAAATGAAG	4260
OY	4262	gttttccaaggnaaatcaagaactctctcaaacctgtgttgtaactctctctagaccacacgttta	4321



|||||  
Db 4261 GTTTTCAGGAATCAAGACTCTTCACACCTGTGTGAACCTCTAGACCCACGCTTA 4320  
Qy 4322 CTGACGCGTACCTCGAATTCACCCAGAGTTGGTGCAGGATTCGCTTAAAGATG 4381  
Db 4321 CTGACTCGCTACTTGCATTCACCCAGAGTTGGTGCAGGATTCGCTTAAAGATG 4380  
Qy 4382 gaggtctgagctcgagagcagagactctactga 4417  
Db 4381 GAGGTCTGGGCTGGAGCAGACAGACTCTACTGA 4416

## RESULT 2

US-08-683-839B-2  
Sequence 2, Application US/08683839B  
Patent No. 5744326

## GENERAL INFORMATION:

APPLICANT: ILL, Charles . R. et al.  
TITLE OF INVENTION: Use of Viral Cis-Acting Post-Transcriptional  
TITLE OF INVENTION: Regulatory Sequences To Increase Expression of  
TITLE OF INVENTION: Introns Genes Containing Near-Consensus Splice Sites  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 60 State Street, suite 510  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109-1875

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/683,839B  
FILING DATE: 11-MARCH-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:

## FILING DATE:

ATTORNEY/AGENT INFORMATION:  
NAME: Remillard, Jane E.  
REGISTRATION NUMBER: 38,872  
REFERENCE/DOCKET NUMBER: TTI-138  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)227-7400  
TELEFAX: (617)227-5941

## INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:  
LENGTH: 9354 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 2965..7378  
US-08-683-839B-2

Query Match 99.8%; Score 4407.4; DB 1; Length 9354;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 441; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 catgcaaatagagctctcaacgtctcttctgtgctcttttgcatcttgcttaagtc 60  
Db 2964 CATGGTTTATGAGCTCTCCACCTGCTTCTTGTGTGCTTTTGGATTCTGCTTTAGTGC 3023  
Qy 61 caccagaataactactcctggtgagtgagacgtgcatgagactatgagactatgcaagtgtact 120  
Db 3024 CACGAGAAGACTACTCTGGGTGAGTGGAAGTGTCTATGGGACTATATGCAAAAGTGTCT 3083

Qy 121 cagtgagctgctgtgagcgaagattctcctagagtgccaaatcttccaltcaa 180  
Db 3084 CGGTGACCTGCTGTGGAGCCAGGATTTCTCTAGAGTGCAGAAATCTTTTCCATTCAA 3143  
Qy 181 cactcagtcgtgtacaaaagactcgtttgtagaattcaagttcactttcaact 240  
Db 3144 CACCTCAGTGTGTACAAAAGACTGTTGTGTGAATTCAGGTTTCACTTTTCAACAT 3203  
Qy 241 cgcctaaagcgaagcaccctgagatggtctgctagtgctccatccaaagctgagttta 300  
Db 3204 CGTTAAGCCAAAGGCCACCTGAGATGGGTCTGTAGTGTCTTACCATCAAGCTGAGTTTA 3263  
Qy 301 tgatacagtggtcaatctaaccttaagaacatggtcctccatctgtcaagtcctatgct 360  
Db 3264 TGATACAGTGTGTATACCTTAAAGAACATGGCTTCCATCTGTGTAAGTCTTATCTGT 3323  
Qy 361 tgggtatcctactcggaaagctctcgtgagagcttgaaatgatgatcgaagcagtgataag 420  
Db 3324 TGGTGTATCCTACTGGAAGCTCTGAGGAGCTGAAATGATGATCAGCCAGTAAAG 3383  
Qy 421 ggaagaagaatgataaagcttccctggttggaagccatataatgctgagagtcct 480  
Db 3384 GGAAGAAAGAAATGATATAAGTCTTCCCTGGTGAAGCCATACATATGTCTGGCAGGTCT 3443  
Qy 481 gaaagaatggtcccaatggtcctctgaccccaatggtccttaactatcatcttctca 540  
Db 3444 GAAAGAAATGTTCAATGAGCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTG 3503  
Qy 541 tgtgagacctgtgtaaaagacttgaaatcagccctcaatgagagcctctactgtatgataa 600  
Db 3504 TGTGACCTGCTTAAGAACTGTAATTCAGGCTCATATGGAGCCCTCATGATGTAGAGA 3563  
Qy 601 aggaagctggtgccaaggaagaaagacacagacctgtgacaatcttataactcttctgt 660  
Db 3564 AGGAGTCTGGCCAAAGGAAAGAAAGACAGACCTTGCAAAATTAATATCTTTTCTGT 3623  
Qy 661 atttgatgaaggaagattggtgacctagaacaaagaactccttcatgtagtaggga 720  
Db 3624 ATTTGATGAAGGGGAAAGTTGGCACTCAGAAACAAACAACTCTTATGAGATGGGA 3683  
Qy 721 tgcgtcactgctgctggtggtcgtgaaatgacacagctgaatggtatgataaaggtc 780  
Db 3684 TGTGCAATCTGCTGGGCTGCGCTTAAGATGACACAGTAAATGTTATGTAACAGGTC 3743  
Qy 781 tctccagcttgatgtgagtcacagaagaaatcagtcctatgtgcatgtgagttggaatgg 840  
Db 3744 TCTGCCAGTCTGTGATGTGAGTGCACAGGAATCATGTTATGGCAATGTGATGGATGG 3803  
Qy 841 caccactcctgaagtgactcaatcatctcctogaaggtcacacattctctgtgaggaaca 900  
Db 3804 CACCACCTCTGAAGTGCACATCAATATCTCGAAGGTGCACACATTTCTGTGAGGAACA 3863  
Qy 901 tcgccaaggtcctctggaatctgcgaataacttcttaactggtcgaacactctgtgt 960  
Db 3864 TCGCCAGGCGCTTGGAAATCTGCCAATPACTTCTTATGCTCAACACCTTGTGAT 3923  
Qy 961 ggaacctggaagtttctactggtttgtcatatctcttccacaacaatgatgtgcatgga 1020  
Db 3924 GAGCCTTGAGACATTTTACTGTTTGTATCTCTTCCACACATGATGTGATGGA 3983  
Qy 1021 agcttatgtcaagtagacagctgtccagaggaagcccaactcgaatgaaataatga 1080  
Db 3984 AGCTTATGTCAAAATPAGACAGCTGTCCAGAGGAACCCCACTCGAATGAAATPAGAT 4043  
Qy 1081 agaagcggagactatgatgatctctactatggttctgaaatgagtgtgtgagttga 1140  
Db 4044 AGAAGCGGAAGACTATATGATGATCTTACTGATCTGAAATGATGATGCTGAGTTGA 4103  
Qy 1141 tgaagcaactctcctccttccatccaaatctcgtcagttgccaagaagaagatcctaaac 1200  
Db 4104 TGATGACACACTCTCTCTTCTTATCCAAATTCGCTAGTTGCCAAGACATCTTAAAC 4163  
Qy 1201 ttgggtacattaatgtcgtcgtgaagaggaagactgggactatgctcctttagtctcgc 1260



|||||  
Db 4164 TTGGATACATTTACATTTGCTGCTGAAGAGGAGACTGGACTATGCTCCTACTCCTCG 4223  
QY 1261 ccccgatgacagaagtataaagltcaataattgaaatgagccctcagcgatggtgag 1320  
Db 4224 CCCCAGTGAAGAGATTAATAAAGTCAATATTTGAACAATGGCCCTCAGGGGATTTGGTATG 4283  
QY 1321 gaagtaaaaaagtcggattatgcatcacagatgaaccttaagaatcgtgaagc 1380  
Db 4284 GAAGTCAAAAAAGTCCGATTTATGGCATACAGATGAACCTTTAAGCTCTGTGAAGC 4343  
QY 1381 taticgacatgaatcagaatcttggaccttacttctatgaggaaattggagacacac 1440  
Db 4344 TATTGACATGAATCAGGAATCTTGGACCTTACTTTATGGGAAGTGGACACACACT 4403  
QY 1441 gttgatataattaaagaatcaagaacagacataaactcaccctcaaggaaatcac 1500  
Db 4404 GTTGATTAATTAATTAACAAACCAAGACCATATACTACCTCCACGGAATAC 4463  
QY 1501 taatgctcgtccttgattcaaggagattaccaaaggltgaacatttgaagatt 1560  
Db 4464 TGATGTCCTGCTTTTATTCAAGGAGATTACCAAAAGGTGAACATTTGAAGATTT 4523  
QY 1561 tccaattctgcagagaaatattcaataataatgagagtgactgtagaagatgggac 1620  
Db 4524 TCCAAATTCTGCCAGAGAAATATTCAATATAAATGACAGTACTGTGAAGATGGGC 4583  
QY 1621 aactaaataagatcctcgtgtgctgacccgctatactactagtttgcgttaataaggag 1680  
Db 4584 AACTAATATAGATCCCTGGGTGCTGACCGCTATATCTAGTTTGTAAATATGAGAG 4643  
QY 1681 agatcagcttcaagatcattgagccctcctcactctctacaagaatctgtataata 1740  
Db 4644 AGATCTAGCTTCAAGGACTATTTGGCCCTCTCCTCATCTGCTAATAAAGATCTGTATGTA 4703  
QY 1741 aagaagaaacagataatgtcagacagaagaaatgcatccgttcttctgtatattgata 1800  
Db 4704 AAGAGAAACCAAGATTAATGTACAGACAGAGAAATGTCATCCGTGTTTCTGATTTATGTA 4763  
QY 1801 gaacgaagctgtgtacctacacagaaataaagaagcttctcccaatcagctgagat 1860  
Db 4764 GAACCAACCTGTGTACTCTACAGAGAAATATACAAACCTTCTCCCAATCCAGCTGAGAT 4823  
QY 1861 gaagcttggagatccagagattccaagcctcaacatcatcagacaagataaagtcatt 1920  
Db 4824 GCAGCTTGAAGATCCAGATTTCCAAAGCTTCCAAATCATATGACAGATCAATAGCTATGT 4883  
QY 1921 ttctgaatgttgcagttgttcagttgtttgtcatgagtgagctactgtgtacatctaa 1980  
Db 4884 TTTTGATAGTTGCAAGTGTCAAGTTGTTTGATGAGTGGCATACTGGTACATTTCTAAG 4943  
QY 1981 catltagaagacagactgactccttctctctctctctctctgtataacttcaaacaa 2040  
Db 4944 CATTTGAGACACAGACTGATCTTCTCTCTCTCTCTCTGTGATATACCTTCAACACAA 5003  
QY 2041 aatgtgtctatgaagacacactacactcttcccatctcaggagaaacgtgtctatgac 2100  
Db 5004 AATGGCTATGAAGACACACTACCTATTTCCATTTCTAGAGAAACGTCTTCTCATGTGC 5063  
QY 2101 gatggaanaacccaggtctatgtatcttgggtgtccacaactcagacttgcgaaacagag 2160  
Db 5064 GATGGAANAACCCAGGTCATGATGTTCTGGGTGCCACAACTCAGACTTTCGGAACAGAG 5123  
QY 2161 catgacgccttactgaaaggttctctagttgtgacaagaacactggtgtatataagaga 2220  
Db 5124 CATGACCGCTTACTGAAAGGTTTCTAGTTGTGACAAAGAACACGTGGATTTATTCACAGA 5183  
QY 2221 cagttatgaagatatttcagactactgctgtgaatlaaaaaaatgccaatggaaccaaga 2280  
Db 5184 CAGTTATGAAGATATTTCAGCATACTTCTGTGAGTAAAAAATGCCATTGAAACCAAGAG 5243  
QY 2281 ctctcccaaatctcaagacaccttagcaactaggaanaacaaatlaatgccaaccac 2340  
|||||

Db 5244 CTCTCCAGAAATTTCAAGACACCCCTAGCACTAGCAAAAAAGCAATTTATGCCACCCACC 5303  
QY 2341 agtcttgaagaagccatcaacggaataactcgttactactcttcagltcagatcaagaga 2400  
Db 5304 AGCTTGAAGACCCATCAACGGGAAATTAACCTGACTACTCTTCAGCTCAGATCAAGAGA 5363  
QY 2401 aattgactatgatataccatcatcagttcaaatgaagaagaagatttgaacttataga 2460  
Db 5364 AATTGACTATGATGATATCCATATACAGTTGAATGAAGAAAGAAATTTTGACATTTATGA 5423  
QY 2461 tgaaga tgaanaalcagagcccccagcttccaagaanaacacagacatttataatg 2520  
Db 5424 TGAGATGGAANAATCAGAGCCCCCGCAGCTTTCAAAAGAAAACGACACTATTTATGTC 5483  
QY 2521 tgcagtggagaaggtcttggatataatggatgagatcctcccaatglttcaagaacag 2580  
Db 5484 TGCAGTGGAGAGCTCTGGATTTATGGAGATGACTCCCAATGTCTTAAGAAACAG 5543  
QY 2581 ggtcagagtgagcagtgctccctcagttcaagaagaatggtttccaggaatttactgatg 2640  
Db 5544 GGCTCAGAGTGGCAGTGTCTCCCTCAGTTCAAGAAAGTTGTTTCCAGGAATTTACTGATG 5603  
QY 2641 ctcccttactcagcccttataccgttggagaactaaatgaacatttggactcctgggac 2700  
Db 5604 CTCTTTACTCAGCCCTTATACCGTGAGAACTAATGAACATTTGGAGACTCCTGGGGCC 5663  
QY 2701 atataagaagcagaagtttgaagaataataatcatglttaacttcaagaatcagccctcg 2760  
Db 5664 ATATATGAAGCAGGAAGTTGAAGATATATCATAGTAATCTTTAGAAATCAGGCCCTCG 5723  
QY 2761 tccctatctcttacttactcagcttatttcttataaggagaagatcagaagcagaagaca 2820  
Db 5724 TCCTTATTCCTTTATTTATTTAGCCTTATTTCTTATGAGAGATCAAGCAAGCAAGACAG 5783  
QY 2821 acctagaaaaaacttltgtaagccttaatgaacacaaacttacttltggaagatgcaaca 2880  
Db 5784 ACTAGAAAAAATTTGTCAAGCCTAATGAACCAAACTTACTTTGGAAATGCAACA 5843  
QY 2881 tcaatgtgcaaccacataaagatgttgaactgtgacgtgcaaacgcttggttattctctgagt 2940  
Db 5844 TCATATGGACCCCAATGAAGATGAGTTTGAATGCAAGCAAGCCTGGCTTATTTCTCTATGA 5903  
QY 2941 tgaactggaaaaagatgttgcactcagcgttgaatggacccttctgtgtcgcacactaa 3000  
Db 5904 TGACCTGGAANAAGATGTGCATCAGGCTGATGGACCCCTTCTGTGTCACACTAA 5963  
QY 3001 cacactgaacccctgtcatcagtggaacaagtgaacagtacaggaattgtgctgtgttccac 3060  
Db 5964 CACACTGAACCCCTGCTCATGGGAGACAAAGTGAAGCAATTTGCTGTGTTTTCAC 6023  
QY 3061 catcttltgtatggaccacaaagctgtgtacttcaacttgaataatggaagaacgtgcagggc 3120  
Db 6024 CATCTTTGATGAGACCAAAAGCTGTACTTCACTGAAATATGGAAGAAACGACGGGC 6083  
QY 3121 tccctgcaatcagaatggaagatcccaactttaaagaagaattatcgtctccatcaat 3180  
Db 6084 TCCCTGCAATATCCAAATGGAAGATCCCACTTTTAAAGAAATATTCGCTTCATCAAT 6143  
QY 3181 caatggtctacataatgatacactaccgtgttagtaatggtcagaatcaagaatcag 3240  
Db 6144 CAATGCTACATTAATGATGATACACTACTGCTGTAGTAAATGGCTCAAGATCAAAAGGATTCG 6203  
QY 3241 atgtgtatctgtccagatagggcagcaatgaanaacatccattctattcatctcaagtga 3300  
Db 6204 ATGTATCTGTCTCAGATGGGACGCAATGAANAACATTCATTTCAATTCAGTGAACA 6263  
QY 3301 tgtgttcaactgttgaagaaaaaagagagatataaatggaactgttaacatctcatcag 3360  
Db 6264 TGTGTTCACTGTACGAAAAAAGAGAGATTAATAATGGCACTGTACATCTCATCCAG 6323  
QY 3361 tgttlttgaagcagtgaaatgttaccatccaaagcttgaattgtgcggttgaatgct 3420  
Db 6324 TGTTTTGAAGACAGTGAAGATGTTACATTCAAAGCTGGAATTTGGCGGTGAATGCT 6383  
|||||







|||||  
Db 444 AGGAGAGAAAGATGATATAAGTCTTCCCTGAGGAAGCCATACATATGCTGGCAGTGC 503  
QY 479 ctgaaagaatggtccaatgagcctctgagcccaactgctcttaactatattcttct 538  
Db 504 CTGAAAGAGATGGTCCAAAGGCGCTGACCCACTGTCCTTACTATATCTTTCT 563  
QY 539 catgtgacctggttaaaagacttgaattcagagcctcatcttgagccctactagatgaga 598  
Db 564 CATGTGACCTGGTAAAGACTTGAATTCAGGCTCATTTGAGCCCTACTAGTATGAGA 623  
QY 599 gaaggagctctggccaaggaagaaacagacctgacaaatttactacttctgt 658  
Db 624 GAAGGGAGTGTGGCCAAAGGAAAGACACAGACCTTGACAAATTTACTACTTTTGTCT 683  
QY 659 gttattgataaggaagaaagtgtgacctcagaaacaaagaaccttctgtatgcagatag 718  
Db 684 GTATTGTGATGAAGGAAAGTTTGGCACTCAGAAACAAAGAACTCTTGATGACGAGATAGG 743  
QY 719 gatgctgcatctgtcgggacctgagcctaataatgacacaaatcatatgttatagttaacag 778  
Db 744 GATGCTGCATCTGCTGGGCGCTGGCCTAAATGCACACAGTCAATGGTTATGTAAACAGG 803  
QY 779 tctctgccaagttctgattggaatgcacaaagaaatcagttctatgtgcatgtgattggaat 838  
Db 804 TCTGTGCAGAGTCTGATTGTGATGCCACAGGAATCAGTCTATTGGCATGTGATGGATG 863  
QY 839 ggcacacactcctgaaatgacatcaatattcctcctgaaggtcacaacattctgtggaagac 898  
Db 864 GGACACACCTCTAAATGCACTCAATAATTCCTGAAAGGTACACAAATTTCTTGAGAGAAC 923  
QY 899 catcgccaagcgctcctctggaatctcgccaataacttctctactctgctcacaacaccttg 958  
Db 924 CATCGCAGGCGCTCTTGGAATCTGCGCAATAAATCTTCTTACTGCTCAAAACCTCTTG 963  
QY 959 atggaccttgagacagttctactgtttctgcatatctctccacaacatgattgcatg 1018  
Db 984 ATGACCTTGGACAGTTCCTACTGTTTGTGCATATCTCTCCACCAACATGATGCGCATG 1043  
QY 1019 gaagcttatgtcaaaatgaagcagctgtccagagaaacccaactagaatgaaataat 1078  
Db 1044 GAAGCTTATGTCAAGATGACAGACTGTCCAGAGAAACCCCACTACGAATGAAGAAATAAT 1103  
QY 1079 gaagaagcggaaagactatgataatgatacttactgattctgaaatgagatggtcaggtt 1138  
Db 1104 GAGAAAGCGGAACATATGATGATGATCTTACTGATTCGAATGATGATGGTGGTCAAGTTT 1163  
QY 1139 gatgatgacaactctcctctcttactlccaaatcgctcagttgccaagaagacatcctaa 1198  
Db 1164 GATGATGACAACTCTCTCTCTTATTCCAAAATTCGCTCAGTTGCCAAGAAAGCATCTTAA 1223  
QY 1199 acttggtatcatcaactctgctgctgtaagaggaagactgggactatgctcctctagtcttc 1258  
Db 1224 ACTTGGGTACATTACATTCCTGCTGAAGAGGAGACTGGGACTGATCTCTCTTATGCTCTC 1283  
QY 1259 gcccccagatgacgaagttataaagtcaatattgaaacaatggccctcagcgatgtgt 1318  
Db 1284 GCCCCGATGACGAAGTTRTAAAGTCAATATTGAAACAATGGCCCTCAGCGGATTTGGT 1343  
QY 1319 aggaaglacaaaaaagttccgattatgcatatcacagatgaaaccttlaagactgtgaa 1378  
Db 1344 AGGAATTAACAAAAGTCCGATTTATGCAATACAGATGAACCTTTAAAGACTCTGAA 1403  
QY 1379 gctatcagcatgaatcaagaaatcttgggaccttactcttataggggaagtgtgagacaa 1438  
Db 1404 GCTATTTCAGATGATCAGGAATCTTGGAACCTTTACTTTATGGGGAAGTGTGAGACACA 1463  
QY 1439 ctgttatattatataagaatcaagaacagacacatatacattccactccagaaac 1498  
Db 1464 CTGTGATTTATATTAAAGATCAAGACAGACCATATTAACCTTCACTCCGGAATC 1523  
QY 1499 actgatgtccgtctctgttattccaagagattaccaaaagltgaaacatttgaagat 1558  
|||||

Db 1524 ACTGATGTCCCTCTTGTATTCAGAGATATACCAAAAAGTGTAAACATTTGAAGAT 1583  
QY 1559 ttccaattctgcgaagagaatattccaataataatgacagtgactgtgaagaatgtg 1618  
Db 1584 TTTCCATTTCTGCAGAGAAATATTCAAATATATGACAGTGCATGTAGAAAGTGGG 1643  
QY 1619 ccaactaaatcagatcctctgtctgacccgcatatacctactgttctgtaataatgag 1678  
Db 1644 CCAACTAAATCAATCTCTGGTCTGACCGGCTATTACTCTAGTTTCGTAAATATGGAG 1703  
QY 1679 agagatcctagcttcaagactcatttggccctctccatcctgtctacaagaatctgtagat 1738  
Db 1704 AGAGATCTAGCTTCAGGACTCATTTGGCCCTCTCTCATCTGTACAAAGATCTGTAGAT 1763  
QY 1739 caaagagaaaccaaataatgtcagacaagagaatgtcatcctgttctctgtatttgtat 1798  
Db 1764 CAAGAGGAAACCGATTAATGTGACAAAGAGAAATGTATCTCTGTTTCTGTATTTGAT 1823  
QY 1799 gagaacccaagctgtgtactcctcacagagaatatacaacgcttctcccaatccagcttga 1858  
Db 1824 GAGAACCGAAAGCTGTGATCTCACAGAGAAATATCAACGCTTTTCCCAATTCAGCTGGA 1883  
QY 1859 gtgcaacttgaggaatccagaagttccaagcctccaacatcatgacagacatcaatgtgcat 1918  
Db 1884 GTGACACTTGAGGATCCAGAGTTCGAAGCCTCCACATCATGACAGCATGATGGCTAT 1943  
QY 1919 gtttttgaaagtttcagttgtcagtttgttttgcaatgaggttgacatctgtacattcta 1978  
Db 1944 GTTTTGTATAGTTTGAGTGTGATGATGTTGTTGTCATAGAGTGGCATATCGTCAATTCTA 2003  
QY 1979 agcatggaagcagagactgaactcctctctctctctctctctctctctctctctctctctct 2038  
Db 2004 AGCATTTGAGACACAGACTGACTCTTCTTCTGCTTCTTCTCTGATATATCCTTCAACAC 2063  
QY 2039 aaaaatgctatagaagaacaactacacctatcccatlctcagagagaacatgtctcatatg 2098  
Db 2064 AAAATGCTATGAAGACACACTCACCTTATCCATTTCTCAGAGAAACCTGCTTCATG 2123  
QY 2099 tctgattgaaacccaagttctatgtatctgtgggttgccaacactcaacttctggaacaga 2158  
Db 2124 TCGATGAAACCCAGAGTCTATGATGTGGGTGCGCAACACATCACTTTCGGAACAGA 2183  
QY 2159 ggcataccgctctactgaaggtttctagtctgacaaagacactgtgattatagaa 2218  
Db 2184 GCGATACCCGCTTACTGGAAGTTCCTAGTGTGACAAAGACCTGTGATTTATACGAG 2243  
QY 2219 gacagttatgaagataltccaagactctgtctgagtaaaaaaacaatgcatatgaaaccaaga 2278  
Db 2244 GACAGTTATGAAGATATTTACAGCATCTTGCTGAGTAAAAACAATGCCATTTGAACCAAGA 2303  
QY 2279 agcttctccagaatccaagaacacctagcactaggcaaaagaaatttaatgcacccca 2338  
Db 2304 AGCTTCTCCAGAAATTCAGAGACACCTTACACATGAGCAAAAGCAATTTAATGCCACCCA 2363  
QY 2339 ccagctctgaaacgcatcaacgggaataaactcgtactactcttcaagtcagaatgaag 2398  
Db 2364 CCACTTTGAAGACCATTAACGGGAATTAACCTGACTACTCTTATAGTCAGATCAAGAG 2423  
QY 2399 gaaattgactatgataccaatcagttgaatgaaatgaaagaaagatttgaacattat 2458  
Db 2424 GAAATTGACTATGATATACATATCATGTTGAATGAAGAAAGAAAGATTTTGAATTTAT 2483  
QY 2459 gatgaggaatgaaatcagaagccccgcgaacttcaaaaagaaacagacacatttatt 2518  
Db 2484 GATGAGGATGAATAATAGAGCCCCCGACCTTTCAAAAACAAACAGAACATATTTTAT 2543  
QY 2519 gctgcagtgagagagctctgggattatggatgagatgagttccccaatgttctaaagaaac 2578  
Db 2544 GCTGCGATGAGAGAGCTCTGGGATTTATGGGATGATAGTCTCCCAATATTTCTTAAGAAAC 2603  
QY 2579 agggctcagaatgagtgatctcctcagttccaagaaagtgtttccaagaaattactgat 2638  
Db 2604 AGGGCTCAGAGTGGCAGTGTCCCTCAGTTCAAGAAAGTTGTTTCCAGGAATTTACTGAT 2663  
|||||







Query Match	97.8%;	Score 4319.8;	DB 4;	Length 4999;
Best Local Similarity	99.08;	Pred. No. 0;		
Matches 4373;	Conservative	0;	Mismatches 2;	Indels 42;
				Gaps 1

OY	961	ggacccttggacaggttcttactgtttgttgcaatactcttcccaacaatgatgacatga	10270
Db	1375	ggaccttggacaggttcttactgtttgttgcaatactcttcccaacaatgatgacatga	14344
OY	1021	agctatgtcaaaatgaacagctgtcccaaggagaccaccaactacagaatgaanaataatga	10860
Db	1435	agctaatgtccaagtgagacagctgtcccaaggagaccaccaactacagaatgaanaataatga	14944
OY	1081	agaagcggaaagacatcatgatcatcttaactgtattcttgaatgtgtagtgcgaagtgtga	11440
Db	1495	agaagcggaaagacatcatgatgtactcttaactgtattcttgaatgtgtagtgcgaagtgtga	15554
OY	1141	tgatgtacaactctctctctctttatccaacttcgtctcagttgtccaaagaacatctcaaac	12000
Db	1555	tgatgtacaactctctctctctttatccaacttcgtctcagttgtccaaagaacatctcaaac	16144
OY	1201	ttgggtgcatcttaactgtctgtctgaagaaggagacttggagctatgtcttccttagtctctgc	12660
Db	1615	ttgggtgcatcttaactgtctgtctgaagaaggagacttggagctatgtcttccttagtctctgc	16744
OY	1261	ccccgatgtacagaaagtgtatataaagccaataatttggacaacttgcccttcagcggattgtgag	13200
Db	1675	ccccgatgtacagaaagtgtatataaagccaataatttggacaacttgcccttcagcggattgtgag	17344
OY	1321	gaagttacaaaaaagttccgattttatgtgcatacacagaatgaaccttgaagactgtgaagc	13800
Db	1735	gaagttacaaaaaagttccgattttatgtgcatacacagaatgaaccttgaagactgtgaagc	17944
OY	1381	tattcagcatgtgaatcaggaatctttggaccttacttttgggggaagtgttggagacaact	14400
Db	1795	tattcagcatgtgaatcaggaatctttggaccttacttttgggggaagtgttggagacaact	18544
OY	1441	gttgattatatttaagaatccaagcaagccaacataataacatctccctccaagsgaatcac	15000
Db	1855	gttgattatatttaagaatccaagcaagccaacataataacatctccctccaagsgaatcac	19144
OY	1501	tgatgtccgtctcttgtatattccaagsgaatacccaaaagttgaanaacatttgaagaattt	15600
Db	1915	tgatgtccgtctcttgtatattccaagsgaatacccaaaagttgaanaacatttgaagaattt	19744
OY	1561	ttcaattctgtccaggaagaataattcaaatataatgatgacaagtgtatagaagaatgggccc	16200
Db	1975	ttcaattctgtccaggaagaataattcaaatataatgatgacaagtgtatagaagaatgggccc	20344
OY	1621	aacccaatcagatctctcgtgtgccttgaccggcctatctcttagtttcogttgaatatgtgaag	16800
Db	2035	aacccaatcagatctctcgtgtgccttgaccggcctatctcttagtttcogttgaatatgtgaag	20944
OY	1681	agatctagctctcagactcaatttgccctctctctcatctgtctacaagaagaatctgtatgaca	17400
Db	2095	agatctagctctcagactcaatttgccctctctctcatctgtctacaagaagaatctgtatgaca	21544
OY	1741	aagagggaaaccagataatgtgccaaagaagggaattgtcatcccgttttctgtcatatttgaga	18000
Db	2155	aagagggaaaccagataatgtgccaaagaagggaattgtcatcccgttttctgtcatatttgaga	22144
OY	1801	gaaccggaagctgtggaaccttcacagagaataatacaagctttccccaatccacgctgtgggt	18600
Db	2215	gaaccggaagctgtggaaccttcacagagaataatacaagctttccccaatccacgctgtgggt	22744
OY	1861	ggacgcttgaagatccagagtttccaagcctctccaacatcatgtacaagcatcaataytgctatgt	19200
Db	2275	ggacgcttgaagatccagagtttccaagcctctccaacatcatgtacaagcatcaataytgctatgt	23344
OY	1921	ttttgataagtttgcagttgcagtgttggtttgcatagagtggtgacatctgtatattcttaag	19800
Db	2335	ttttgataagtttgcaagtttgcagtgttggtttgcatagagtggtgacatctgtatattcttaag	23944
OY	1981	catgttgagcaacagatgaaactctctctgtctctctctctctgtgatatcacttccaacaaca	20400
Db	2395	catgttgagcaacagatgaaactctctctgtctctctctctctgtgatatcacttccaacaaca	24544
OY	2041	aatgtgtatagaaagacacatctccatcttcccatctcagagaagaactgtcttcatgtc	21000



|||||  
Db 2455 aatggtctatgaagacaactccacctatccattccattccaggaaactgltcttcagtgc 2514  
OY 2101 gatggaaaaacccaggtctatggaattctgggggtccacaactcagactttcggaaacagg 2160  
Db 2515 gatggaaaaacccaggtctatggaattctgggggtccacaactcagactttcggaaacagg 2514  
OY 2161 catgaccgcttactctaaaggttctcagttgtgacaagaagaactggtattatctacggaga 2220  
Db 2575 catgaccgcttactctaaaggttctcagttgtgacaagaagaactggtattatctacggaga 2614  
OY 2221 cagttatgaagataattcagacatactgtcgtgagttaaaacaatgccaatgaaacagaag 2280  
Db 2635 cagttatgaagataattcagacatactgtcgtgagttaaaacaatgccaatgaaacagaag 2694  
OY 2281 ctctcccccgaattcaagacaccctagcactagcgaagaacatttaatgcccaccacc 2340  
Db 2695 ctctcccccgaattcaagacaccctagcactagcgaagaacatttaatgcccaccacc 2712  
OY 2341 agtcttgaaaacgcatcaacggaataaactcgttactactcttcagtcagatcaagaga 2400  
Db 2713 agtcttgaaaacgcatcaacggaataaactcgttactactcttcagtcagatcaagaga 2772  
OY 2401 aattgactctgatgatccaatcattcagttgaaatggaagaagaattttacatttga 2460  
Db 2773 aattgactctgatgatccaatcattcagttgaaatggaagaagaattttacatttga 2832  
OY 2461 tgaagatgaaatcagaagcccccagcgttccaagaagaacacgaacattttatctgc 2520  
Db 2833 tgaagatgaaatcagaagcccccagcgttccaagaagaacacgaacattttatctgc 2882  
OY 2521 tgcagtgagagaggtctctgagatattggaatgagtaagtcctccacatgtctctaagaacag 2580  
Db 2893 tgcagtgagagaggtctctgagatattggaatgagtaagtcctccacatgtctctaagaacag 2952  
OY 2581 ggctcagagtgagcagtgctccctcagttcgaagaagtggttttccaggaatttactgtg 2640  
Db 2953 ggctcagagtgagcagtgctccctcagttcgaagaagtggttttccaggaatttactgtg 3012  
OY 2641 ccccttactcagcccttaacacgttgagaagaactgaagaactttggagctccctggggcc 2700  
Db 3013 ccccttactcagcccttaacacgttgagaagaactgaagaactttggagctccctggggcc 3072  
OY 2701 atataaagagcagaagttgaaagataatcaltcgttaacttccagaatcagggcctctcg 2760  
Db 3073 atataaagagcagaagttgaaagataatcaltcgttaacttccagaatcagggcctctcg 3132  
OY 2761 tccctattctcttatcttctagccttatttcttatgagaagaatcagaaggcagaagaga 2820  
Db 3133 tccctattctcttatcttctagccttatttcttatgagaagaatcagaaggcagaagaga 3192  
OY 2821 accctagaaaaaacttctgcaagccttaatgaaacaaactacttttggaaagtgcaca 2880  
Db 3193 accctagaaaaaacttctgcaagccttaatgaaacaaactacttttggaaagtgcaca 3252  
OY 2881 tcaatatgcaacccactaaagaatgagtttgaatgcaaaagccttgggtatttctctgagt 2940  
Db 3253 tcaatatgcaacccactaaagaatgagtttgaatgcaaaagccttgggtatttctctgagt 3312  
OY 2941 tgaacctggaaaaaagatgtgtcactcagcctgagtgtgaccccttctgtctgcacactaa 3000  
Db 3313 tgaacctggaaaaaagatgtgtcactcagcctgagtgtgaccccttctgtctgcacactaa 3372  
OY 3001 cagactgaacccctgtcactgaggagacagtgatcagaatcttgcgtttttcac 3060  
Db 3373 cagactgaacccctgtcactgaggagacagtgatcagaatcttgcgtttttcac 3432  
OY 3061 catctcttgaatgagacaaagctgtgactcactctgaaataatggaagaacactgcagggc 3120  
Db 3433 catctcttgaatgagacaaagctgtgactcactctgaaataatggaagaacactgcagggc 3492  
OY 3121 tccctcgcaatatccagatggaagaatccaccatttaagagaatatacgtctccaatgc 3180  
|||||

Db 3493 tccctcgcaatatccagatggaagaatccaccatttaagagaatatacgtctccaatgc 3552  
OY 3181 caatggtctacataatgatacactacactcgtggttagttaaagtcacagatcaagaatcgc 3240  
Db 3553 caatggtctacataatgatacactacactcgtggttagttaaagtcacagatcaagaatcgc 3612  
OY 3241 atggtatctgtcgaagcattgaggcaatgaaacatccatctcatcttcaatgagga 3300  
Db 3613 atggtatctgtcgaagcattgaggcaatgaaacatccatctcatcttcaatgagga 3672  
OY 3301 tgtgttctacgtctgcaaaaaaagagagataaaatggaacgactgttaacatcttaccgg 3360  
Db 3673 tgtgttctacgtctgcaaaaaaagagagataaaatggaacgactgttaacatcttaccgg 3732  
OY 3361 tgtttttgacacagtggaatgtttacatccaaagcttggaatttggcgggtggaatgcct 3420  
Db 3733 tgtttttgacacagtggaatgtttacatccaaagcttggaatttggcgggtggaatgcct 3792  
OY 3421 tattggcagatctacatgctcgtggatgagacacttttctgtgtgtacagaataatgtg 3480  
Db 3793 tattggcagatctacatgctcgtggatgagacacttttctgtgtgtacagaataatgtg 3852  
OY 3481 tcaagctcccttgggaatggtctcttgacaacatgaagaatttcaagatacagctttcagg 3540  
Db 3853 tcaagctcccttgggaatggtctcttgacaacatgaagaatttcaagatacagctttcagg 3912  
OY 3541 acatatgagcaatgagcccccagacgttgccagactctcatctccggaatcaatcagtc 3600  
Db 3913 acatatgagcaatgagcccccagacgttgccagactctcatctccggaatcaatcagtc 3972  
OY 3601 ctggagacacaaagagaccccttctctggaatcgaagtgtgaatctgttggccaatgattat 3660  
Db 3973 ctggagacacaaagagaccccttctctggaatcgaagtgtgaatctgttggccaatgattat 4032  
OY 3661 tcaagcatcaaaacccaaggtgtcccgctcagaagaagttctccagcttataatcttcagtt 3720  
Db 4033 tcaagcatcaaaacccaaggtgtcccgctcagaagaagttctccagcttataatcttcagtt 4092  
OY 3721 tatcatcatgtatagttcttatatggaagaagtgagcagactatcagaagaattccactg 3780  
Db 4093 tatcatcatgtatagttcttatatggaagaagtgagcagactatcagaagaattccactg 4152  
OY 3781 aaacctaaatggtctctcttggcaatgtggaatcattctgggaataaaacacataattttaa 3840  
Db 4153 aaacctaaatggtctctcttggcaatgtggaatcattctgggaataaaacacataattttaa 4212  
OY 3841 cccctcaatattgtctcgtatataatccgtttgcaacccaactcatatagatctgcagacc 3900  
Db 4213 cccctcaatattgtctcgtatataatccgtttgcaacccaactcatatagatctgcagacc 4272  
OY 3901 tcttcgcatggaatgtgatggcgtgtgatttaataatgttcagacatgccaattgggaatgga 3960  
Db 4273 tcttcgcatggaatgtgatggcgtgtgatttaataatgttcagacatgccaattgggaatgga 4332  
OY 3961 gagtaagaacatatcagaatgcacagattactgtcttcaacttaacataatgtttgc 4020  
Db 4333 gagtaagaacatatcagaatgcacagattactgtcttcaacttaacataatgtttgc 4392  
OY 4021 caactgtctccttcaaaaactgagcttcaacctcgaagggaaggaatgagcttgcagacc 4080  
Db 4393 caactgtctccttcaaaaactgagcttcaacctcgaagggaaggaatgagcttgcagacc 4452  
OY 4081 tcaagtgaaatcaacaaaagatggtctgcaagtgtgacttccagaagaacataagaatcag 4140  
Db 4453 tcaagtgaaatcaacaaaagatggtctgcaagtgtgacttccagaagaacataagaatcag 4512  
OY 4141 aggaatgaactactcagggagtgaaatctcgtcttaccagaatgtatgtgaagagttcct 4200  
Db 4513 aggaatgaactactcagggagtgaaatctcgtcttaccagaatgtatgtgaagagttcct 4572  
OY 4201 catctccagagtcagaatgagcacaatcagtgagactcttcttttcgaatgagcagaagtaa 4260  
Db 4573 catctccagagtcagaatgagcacaatcagtgagactcttcttttcgaatgagcagaagtaa 4632  
|||||



Oy	4261	ggttttttcagggaataacaagactccctcaacactgtggtggaactctctagaccacggt	4320
Db	4633	ggttttttcagggaataacaagactccctcaacactgtggtggaactctctagaccacggt	4692
Oy	4321	actgactcgctactccttcggaattcacccccagagtttggtgtcaccaagattgacctagagat	4380
Db	4693	actgactcgctactccttcggaattcacccccagagtttggtgtcaccaagattgacctagagat	4752
Oy	4381	ggaggttctgggctgcgagacacaggaactctactgga	4417
Db	4753	ggaggttctgggctgcgagacacaggaactctactgga	4789

RESULT 5

US-09-364-862-14

Sequence 14, Application US/09364862

Patent No. 6221349

GENERAL INFORMATION:

APPLICANT: Couto, Linda B.

APPLICANT: Colosi, Peter C.

TITLE OF INVENTION: ADENO-ASSOCIATED VECTORS FOR EXPRESSION OF FACTOR VIII

TITLE OF INVENTION: BY TARGET

TITLE OF INVENTION: CELLS

FILE REFERENCE: AVIGEN-03743

CURRENT APPLICATION NUMBER: US/09/364,862

CURRENT FILING DATE: 1999-07-30

EARLIER APPLICATION NUMBER: 60/125,974

EARLIER FILING DATE: 1999-03-24

EARLIER APPLICATION NUMBER: 60/104,994

EARLIER FILING DATE: 1998-10-20

NUMBER OF SEQ ID NOS: 14

SOFTWARE: Patentin Ver. 2.0

SEQ ID NO 14

LENGTH: 4999

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: Synthetic

US-09-364-862-14

Query Match	Similarity	99.8%	Score 4319.8	DB 4	Length 4999
Matches 4373	Conservative	0	Mismatches	2	Indels 42
					Gaps 1
QY 1	catgaataatagagctctccacactgtctctctctgtgaccttggcgatctgcttaatgac				60
Db 415	catgaataatagagctctccacactgtctctctctgtgaccttggcgatctgcttaatgac				60
QY 61	caccgaagaatactactcctcgtgtgcagctggaactcgtcaatggagctatataatgaagtac				120
Db 475	caccgaagaatactactcctcgtgtgcagctggaactcgtcaatggagctatataatgaagtac				120
QY 121	cgtgtagctgcctcgtggagcgcgaagaatttcctcctatagatgagccaaaatctttccatcaa				180
Db 535	cgtgtagctgcctcgtggagcgcgaagaatttcctcctatagatgagccaaaatctttccatcaa				180
QY 181	cacctcagtcgtatcaaaaagaactcgttttgtatgaattcacgggttcacctttccaat				240
Db 595	cacctcagtcgtatcaaaaagaactcgttttgtatgaattcacgggttcacctttccaat				240
QY 241	cgttaagcccaagagccacacctggaatgggtctgtttaagttcctaccacccagcgtgaagtta				300
Db 655	cgttaagcccaagagccacacctggaatgggtctgtttaagttcctaccacccagcgtgaagtta				300
QY 301	tgatatacagttggtatcatcaacttaagaacaatggcttcccatccctgttaagcttcatgtctg				360
Db 715	tgatatacagttggtatcatcaacttaagaacaatggcttcccatccctgtgtcatgtctcatgtctg				360
QY 361	tgtgtatctactactgtgaagaagcttcgtgagggagactgaatatgtatgtatcagaacagttcaag				420
Db 775	tgtgtatctactactgtgaagaagcttcgtgagggagactgaatatgtatgtatcagaacagttcaag				420

OY	421	ggaagaaagaaatgaataaactcttcccttggatgaagacataatbtgtctggcaagtcct	480
Db	835	ggagaaagaaagaaatgaataaagctcttcccttggatgaagacataatbtgtctggcaagtcct	894
OY	481	gaaagaaatgtgtccaaatgtgctcttgaccccaactgtgtcccttactactacatacttctcca	540
Db	895	gaaagaaatgtgtccaaatgtgctcttgaccccaactgtgtcccttactactacatacttctcca	954
OY	541	ttgtgacctgtgtaaagaacttgatattccagtcctcatctggagccctactactgtatgaagga	600
Db	955	ttgtgacctgtgtaaagaacttgatattccagtcctcatctggagccctactactgtatgaagga	1014
OY	601	agggagctctggcccaaggaagaaagacacagacactgtgcacaatttatactactctttgtcgt	660
Db	1015	agggagctctgtcccaaggaagaaagacacagacactgtgcacaatttatactactctttgtcgt	1074
OY	661	atttgatgaagggaaagaactgtgtgcactcagaacaaagaacatcccttgatgtcaggaatagga	720
Db	1075	atttgatgaagggaaagaactgtgtgcactcagaacaaagaacatcccttgatgtcaggaatagga	1134
OY	721	ttgtgcacatctgtctgtggccctgtgccttaaaatgtgcacaaatgtcatatgtttatgtaaagaatc	780
Db	1135	ttgtgcacatctgtctgtggccctgtgccttaaaatgtgcacaaatgtcatatgtttatgtaaagaatc	1194
OY	781	tcctgcacagctctgtatttgatgtccacaggaataatcgtatttgatgtatgtatgtgaatgtg	840
Db	1195	tcctgcacagctctgtatttgatgtccacaggaataatcgtatttgatgtatgtatgtgaatgtg	1254
OY	841	caccaactctctgaaagtgtcactccaataatctctctgaaggttcacacaaattctgtgtgaagaacca	900
Db	1255	caccaactctctgaaagtgtcactccaataatctctctgaaggttcacacaaattctgtgtgaagaacca	1314
OY	901	tcgcacagagcgctcctgttgaaatctgcgcacaataattctctactgtgtccacaacactgtgat	960
Db	1315	tcgcacagagcgctcctgttgaaatctgcgcacaataattctctactgtgtccacaacactgtgat	1374
OY	961	ggacctctgtgacagttcttactgttctgtcatalctctctcccaacaatgatatgtgcagatga	1020
Db	1375	ggacctctgtgacagttcttactgttctgtcatalctctctcccaacaatgatatgtgcagatga	1434
OY	1021	agcttatgtccaagaatgaagacagctgtccagaaggaaaccccaactcagatgtgaanaaatgaatga	1080
Db	1435	agcttatgtccaagaatgaagacagctgtccagaaggaaaccccaactcagatgtgaanaaatgaatga	1494
OY	1081	agaagcggaaagactatgatgatgtactctcgtattcttgaaatgtgaatgtgtgtcgaagttctga	1140
Db	1495	agaagcggaaagactatgatgatgtactctcgtattcttgaaatgtgaatgtgtgtcgaagttctga	1554
OY	1141	tgatgtacaaactctccctctcttattccaagaatgtcgtatgttgcgaagaagacatctctaaaac	1200
Db	1555	tgatgtacaaactctccctctcttattccaagaatgtcgtatgttgcgaagaagacatctctaaaac	1614
OY	1201	ttgttgatcatataactgtctctctgaaagaagagacactgtgcactatgtctccctttagtctctgc	1260
Db	1615	ttgttgatcatataactgtctctctgaaagaagagacactgtgcactatgtctccctttagtctctgc	1674
OY	1261	cccccgatgcagagaagtataaagaatgcataatttgaaacaaatgtgcctctcagcggatgtgtag	1320
Db	1675	cccccgatgcagagaagtataaagaatgcataatttgaaacaaatgtgcctctcagcggatgtgtag	1734
OY	1321	gaagtataaaaaagtctcgatatttatgtgcataaccagatgaacactttaagactcgtgaagc	1380
Db	1735	gaagtataaaaaagtctcgatatttatgtgcataaccagatgaacactttaagactcgtgaagc	1794
OY	1381	tattccagcatgtaaatcaggaatctcttggaacttactattcttgaggagaagtcttggaacaacct	1440
Db	1795	tattccagcatgtaaatcaggaatctcttggaacttactattcttgaggagaagtcttggaacaacct	1854
OY	1441	gttgtatataatttaagaatacgaacgaagacacataataactctactccctcagcgaatactac	1500
Db	1855	gttgtatataatttaagaatacgaacgaagacacataataactctactccctcagcgaatactac	1914



QY 1501 tgaatgctgcctcttgatctcaagaagattaccaaaggtgtaaacatttgaagatctt 1560  
| | | | |  
Db 1915 tgaatgctgcctcttgatctcaagaagattaccaaaggtgtaaacatttgaagatctt 1974  
| | | | |  
QY 1561 tccaattctgcagggaataatccaataataatgaacagtgactgtgaagaagtggcc 1620  
| | | | |  
Db 1975 tccaattctgcagggaataatccaataataatgaacagtgactgtgaagaagtggcc 2034  
| | | | |  
QY 1621 aactaaatcagatcctgtgctgcagccgctatactctagttctgttaataatgaagaag 1680  
| | | | |  
Db 2035 aactaaatcagatcctgtgctgcagccgctatactctagttctgttaataatgaagaag 2094  
| | | | |  
QY 1681 agactagcttcaagactcattggccctccctccatctgtctacaagaatctgtatgaca 1740  
| | | | |  
Db 2095 agactagcttcaagactcattggccctccctccatctgtctacaagaatctgtatgaca 2154  
| | | | |  
QY 1741 aagaagaaacagataatgtgcagaagaagatgtacatcctgtttctgtatattgata 1800  
| | | | |  
Db 2155 aagaagaaacagataatgtgcagaagaagatgtacatcctgtttctgtatattgata 2214  
| | | | |  
QY 1801 gaaccgaagctgtgtaacctcaagaagaatatacaacgctttccccaatccagctgtgagt 1860  
| | | | |  
Db 2215 gaaccgaagctgtgtaacctcaagaagaatatacaacgctttccccaatccagctgtgagt 2274  
| | | | |  
QY 1861 gcaagcttgagatccagaagttccaagcctccaacatcatgacagatcaatgtgctatgt 1920  
| | | | |  
Db 2275 gcaagcttgagatccagaagttccaagcctccaacatcatgacagatcaatgtgctatgt 2334  
| | | | |  
QY 1921 ttttgatagtttcagatgtgcagttgttgatgaggtgacatggtatcttcaag 1980  
| | | | |  
Db 2335 ttttgatagtttcagatgtgcagttgttgatgaggtgacatggtatcttcaag 2394  
| | | | |  
QY 1981 catgtgagcaacagactgaactccttctgtctctctctctctctctctctctctctctct 2040  
| | | | |  
Db 2395 catgtgagcaacagactgaactccttctgtctctctctctctctctctctctctctctct 2454  
| | | | |  
QY 2041 aatgtgtctatgaagaacacactcaacctatcccatctctcagagagaactgtctcatgtc 2100  
| | | | |  
Db 2455 aatgtgtctatgaagaacacactcaacctatcccatctctcagagagaactgtctcatgtc 2514  
| | | | |  
QY 2101 gatggaaaaacccaggtgtatgtatctctgggtgtgcacaacacagactcttgagaacagag 2160  
| | | | |  
Db 2515 gatggaaaaacccaggtgtatgtatctctgggtgtgcacaacacagactcttgagaacagag 2574  
| | | | |  
QY 2161 catgacgcgcttaactggaagtttctagttgtgacaagaacactgtgatatltacgagga 2220  
| | | | |  
Db 2575 catgacgcgcttaactggaagtttctagttgtgacaagaacactgtgatatltacgagga 2634  
| | | | |  
QY 2221 cagttatgaagataatctcagacatactgtcgtgagtataaaacaatgcatltgaaccaagaag 2280  
| | | | |  
Db 2635 cagttatgaagataatctcagacatactgtcgtgagtataaaacaatgcatltgaaccaagaag 2694  
| | | | |  
QY 2281 ctctccccaagaattcaagacacacctagcacttaaggcaaaagcaatttaatgcccaccgac 2340  
| | | | |  
Db 2695 ctctccccaagaattcaagacacacctagcacttaaggcaaaagcaatttaatgcccaccgac 2712  
| | | | |  
QY 2341 agtcttgaaaaacgcatcaacaggggaataaactcgtactactctctcagtcagatcaagaaga 2400  
| | | | |  
Db 2713 agtcttgaaaaacgcatcaacaggggaataaactcgtactactctctcagtcagatcaagaaga 2772  
| | | | |  
QY 2401 aattgactatgatatcaacatatacagtgtgaatgaagaagaagaatttgaacattatga 2460  
| | | | |  
Db 2773 aattgactatgatatcaacatatacagtgtgaatgaagaagaagaatttgaacattatga 2832  
| | | | |  
QY 2461 tgaagatgaanaatcagagcccgccgagcttccaanaagaacaagacaactatttatgctc 2520  
| | | | |  
Db 2833 tgaagatgaanaatcagagcccgccgagcttccaanaagaacaagacaactatttatgctc 2892  
| | | | |  
QY 2521 tgcagtgagagagctcttgagatatltgagatgagtagctccccaatgttcttaagaanaag 2580  
| | | | |  
Db 2893 tgcagtgagagagctcttgagatatltgagatgagtagctccccaatgttcttaagaanaag 2952  
| | | | |  
QY 2581 ggcctcagaagtgtgcagtgctccctcagatctcaagaagaagtgttcttccaagaatttactgtg 2640  
| | | | |

Db 2953 ggcctcagaagtgtgcagtgctccctcagttccaagaagaagtgttcttccaagaatttactgtg 3012  
| | | | |  
QY 2641 ctctcttactcagaccctataccgttgagagaactaaagaacatttggactctggggcc 2700  
| | | | |  
Db 3013 ctctcttactcagaccctataccgttgagagaactaaagaacatttggactctggggcc 3072  
| | | | |  
QY 2701 atataaagagcagaagtgtgaagataataatcatgtgaacttccaagaatacagccctcg 2760  
| | | | |  
Db 3073 atataaagagcagaagtgtgaagataataatcatgtgaacttccaagaatacagccctcg 3132  
| | | | |  
QY 2761 tccctattcctctatctctgagccttattctctatgaagaatcagaagcagaagcaga 2820  
| | | | |  
Db 3133 tccctattcctctatctctgagccttattctctatgaagaatcagaagcagaagcaga 3192  
| | | | |  
QY 2821 accatagaanaaacttgcagccttaagctaaagcaacaaacttacttcttgaagaatgaaaca 2880  
| | | | |  
Db 3193 accatagaanaaacttgcagccttaagctaaagcaacaaacttacttcttgaagaatgaaaca 3252  
| | | | |  
QY 2881 tcatatgcaaccaccaataagaatgagttgactgcagaagcctgggcttattctctgtatgt 2940  
| | | | |  
Db 3253 tcatatgcaaccaccaataagaatgagttgactgcagaagcctgggcttattctctgtatgt 3312  
| | | | |  
QY 2941 tgaacctggaaaaaagaatgtgcaactcaggtgattggaacccctctgtgtgccaaactaa 3000  
| | | | |  
Db 3313 tgaacctggaaaaaagaatgtgcaactcaggtgattggaacccctctgtgtgccaaactaa 3372  
| | | | |  
QY 3001 cacactgaacccctgtcatctatggagaacaggtgacagttacagaagtctgtctgttttccac 3060  
| | | | |  
Db 3373 cacactgaacccctgtcatctatggagaacaggtgacagttacagaagtctgtctgttttccac 3432  
| | | | |  
QY 3061 catcttctgatatgagaccaaagcgtgttactcactctgaaataatgtgaagaagaactgtcagggc 3120  
| | | | |  
Db 3433 catcttctgatatgagaccaaagcgtgttactcactctgaaataatgtgaagaagaactgtcagggc 3492  
| | | | |  
QY 3121 tccctgtcaatatccagatgtgaagaatcccaactttaaagagaattatcgcttccatgcaat 3180  
| | | | |  
Db 3493 tccctgtcaatatccagatgtgaagaatcccaactttaaagagaattatcgcttccatgcaat 3552  
| | | | |  
QY 3181 caatggtctacataatgtgatacactacactgtgttagtaatgtggtccagagatcaagaagtttg 3240  
| | | | |  
Db 3553 caatggtctacataatgtgatacactacactgtgttagtaatgtggtccagagatcaagaagtttg 3612  
| | | | |  
QY 3241 atggatctgtctcagcaatgggcagcaatgaaacatccatctatcatcttcaagtggaca 3300  
| | | | |  
Db 3613 atggatctgtctcagcaatgggcagcaatgaaacatccatctatcatcttcaagtggaca 3672  
| | | | |  
QY 3301 tgtgttcaactgttacgaaaaaagaagagatataaatgtgcactgttacaactctataccag 3360  
| | | | |  
Db 3673 tgtgttcaactgttacgaaaaaagaagagatataaatgtgcactgttacaactctataccag 3732  
| | | | |  
QY 3361 tgtttttgagacagttgaaaaatgttacaacatccaagctggaatttggcggtggaatgctc 3420  
| | | | |  
Db 3733 tgtttttgagacagttgaaaaatgttacaacatccaagctggaatttggcggtggaatgctc 3792  
| | | | |  
QY 3421 tattgycgagatctacataatgtctggatgtgacacacttctctgtgtgtaagaacaataatgt 3480  
| | | | |  
Db 3793 tattgycgagatctacataatgtctggatgtgacacacttctctgtgtgtaagaacaataatgt 3852  
| | | | |  
QY 3481 tccagacccccctgggaatgtgctctctggacacatataagaatttccaataacagcttcag 3540  
| | | | |  
Db 3853 tccagacccccctgggaatgtgctctctggacacatataagaatttccaataacagcttcag 3912  
| | | | |  
QY 3541 acaatatgagacagttggcccccaagcctgggcagacacttcaatctccggatccaataatgc 3600  
| | | | |  
Db 3913 acaatatgagacagttggcccccaagcctgggcagacacttcaatctccggatccaataatgc 3972  
| | | | |  
QY 3601 ctggaagcacaagaagcccttctcttgatccaaggtgagatctgtgttgccaccaatgatat 3660  
| | | | |  
Db 3973 ctggaagcacaagaagcccttctcttgatccaaggtgagatctgtgttgccaccaatgatat 4032  
| | | | |  
QY 3661 tcaacgcatcaagaaccagaggtgcccgtcagaagaattctccagcctctacatctctcagtt 3720  
| | | | |



```
Db 4033 tcacgagcatcaagaccagggtgcccgttcagaaagtctccagacctctacatctcagtt 4092
QY 3721 tatcatcatgtatagctctgatgtggaagatggtgcagactatcgggaattccactg 3780
Db 4093 tatcatcaatgataagctctgataagggaagatggtgcagactatcgggaattccactg 4152
QY 3781 aaccttaatgtctctctcttgccaatgtgatacatctggatataaacacatacttttaa 3840
Db 4153 aaccttaatgtctctctcttgccaatgtgatacatctggatataaacacatacttttaa 4212
QY 3841 cccctcaatattgtctcgaacacacacacacacacacacacacacacacacacacacac 3900
Db 4213 cccctcaatattgtctcgaacacacacacacacacacacacacacacacacacacacac 4272
QY 3901 tcttcgcatggaatgtatgtggtgtgtatataatgtgagaatgtccatgtggatgga 3960
Db 4273 tcttcgcatggaatgtatgtggtgtgtatataatgtgagaatgtccatgtggatgga 4332
QY 3961 gagtaagacataatcagatcagacagatcagatcagatcagatcagatcagatcagatc 4020
Db 4333 gagtaagacataatcagatcagacagatcagatcagatcagatcagatcagatcagatc 4392
QY 4021 caccctgtctctctcacaagctcgtcactcactcactcactcactcactcactcactc 4080
Db 4393 caccctgtctctctcacaagctcgtcactcactcactcactcactcactcactcactcact 4452
QY 4081 tcaggtgataatcacaagaagatggtgtcagaagatggtgtcagaagatggtgtcagaag 4140
Db 4453 tcaggtgataatcacaagaagatggtgtcagaagatggtgtcagaagatggtgtcagaag 4512
QY 4141 aggaagtaactactcagaaggaatcactcgtctaccagatcagatcagatcagatcag 4200
Db 4513 aggaagtaactactcagaaggaatcactcgtctaccagatcagatcagatcagatcagat 4572
QY 4201 catctcagacagatcagaatggtcagatcagatcagatcagatcagatcagatcagat 4260
Db 4573 catctcagacagatcagaatggtcagatcagatcagatcagatcagatcagatcagat 4632
QY 4261 ggttttcagaggaatcagaagctcctcactcactcactcactcactcactcactcact 4320
Db 4633 ggttttcagaggaatcagaagctcctcactcactcactcactcactcactcactcact 4692
QY 4321 acgagactcgtactcctcgaatcaccacacacacacacacacacacacacacacacac 4380
Db 4693 acgagactcgtactcctcgaatcaccacacacacacacacacacacacacacacacac 4752
QY 4381 ggaagttctggtgctgcagagcagacagacactctactga 4417
Db 4753 ggaagttctggtgctgcagagcagacagacactctactga 4789

RESULT 6
US-09-470-618-13
; Sequence 13, Application US/09470618
; Patent No. 6200560
; GENERAL INFORMATION:
; APPLICANT: Colosi, Peter C.
; APPLICANT: Colosi, Peter C.
; TITLE OF INVENTION: Adeno-Associated Vectors for Expression of Factor VIII
; FILE REFERENCE: Avigen-04082
; CURRENT APPLICATION NUMBER: US/09/470,618
; EARLIER APPLICATION NUMBER: 09/364,862
; EARLIER FILING DATE: 1999-07-30
; EARLIER APPLICATION NUMBER: 60/125,974
; EARLIER FILING DATE: 1999-03-24
; EARLIER APPLICATION NUMBER: 60/104,994
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 11933
```

```
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-470-618-13

Query Match
Best Local Similarity 98.2%; Pred. No. 0;
Matches 4338; Conservative 0; Mismatches 1; Indels 78; Gaps 1;

QY 1 catgcaatagaagctctccacgtctctctctctctctctctctctctctctctctctctct 60
Db 426 catgcaatagaagctctccacgtctctctctctctctctctctctctctctctctctctct 485
QY 61 caccagaagatactactcctggtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 120
Db 486 caccagaagatactactcctggtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 545
QY 121 cgttgagctggtcgtgtgagcagaagattctctcctagagtgagcagaagattctctcct 180
Db 546 cgttgagctggtcgtgtgagcagaagattctctcctagagtgagcagaagattctctcct 605
QY 181 caccctagctcgtgtgacaaaagactctgtgtgagatcagatcagatcagatcagatcagat 240
Db 606 caccctagctcgtgtgacaaaagactctgtgtgagatcagatcagatcagatcagatcagat 665
QY 241 cgtctaaagcagaagcaccctggtgagtggtctgagtgagtgagtgagtgagtgagtgag 300
Db 666 cgtctaaagcagaagcaccctggtgagtggtctgagtgagtgagtgagtgagtgagtgag 725
QY 301 tgataagtggtcatcactaagaaacatggtctccactcctcgtcagctcagctcagctcag 360
Db 726 tgataagtggtcatcactaagaaacatggtctccactcctcgtcagctcagctcagctcag 785
QY 361 tggatgatactactcagaaagctctcagagagagctcagatcagatcagatcagatcagat 420
Db 786 tggatgatactactcagaaagctctcagagagagctcagatcagatcagatcagatcagat 845
QY 421 ggaagaagaagatgataaagctctcctcgtgtggaagcacaataatgctcgtcagctcct 480
Db 846 ggaagaagaagatgataaagctctcctcgtgtggaagcacaataatgctcgtcagctcct 905
QY 481 gaaagaagaatggtcagatggtcctcagacacgtcgtcactactactactactactact 540
Db 906 gaaagaagaatggtcagatggtcctcagacacgtcgtcactactactactactactactact 965
QY 541 tgtgacactggttaaaagactgaattcagagcctcattgagcctcactactactactactact 600
Db 966 tgtgacactggttaaaagactgaattcagagcctcattgagcctcactactactactactact 1025
QY 601 agggagctctgccaagaaagaacacagacactgacacaaattatactactcttctgtct 660
Db 1026 agggagctctgccaagaaagaacacagacactgacacaaattatactactcttctgtct 1085
QY 661 atttgatgaaggaaagattggtcactcagaagaacaaagactcctgtagagagatgagga 720
Db 1086 atttgatgaaggaaagattggtcactcagaagaacaaagactcctgtagagagatgagga 1145
QY 721 tctgcatctgctcgtggtcgtcctaaatgacacagcgaatggtgtaaacagctc 780
Db 1146 tctgcatctgctcgtggtcgtcctaaatgacacagcgaatggtgtaaacagctc 1205
QY 781 tctgcaagctcgtatgtagtccacagaaatcagatctatggtcagatgtagtgaatgag 840
Db 1206 tctgcaagctcgtatgtagtccacagaaatcagatctatggtcagatgtagtgaatgag 1265
QY 841 caccactcctgaagtgacatcaatattcctggaagtgacacatctctgtgaagaaca 900
Db 1266 caccactcctgaagtgacatcaatattcctggaagtgacacatctctgtgaagaaca 1325
QY 901 tcgcagagctccttggaatctcgccaataactctcactactgctcaaacactctgag 960
```



Dh 1326 tgcgcagcgctcttgcgaaactctgcgaataacttcttacttgcgtccaaacactcttgat 1385  
Qy 961 ggaaccttgagcaagtcttcaactgtttgttcatactcttccacacaacatgatagtgcata 1020  
Dh 1386 ggaaccttgagcaagtcttcaactgtttgttcatactcttccacacaacatgatagtgcata 1445  
Qy 1021 agcttatgtcaaaagttagacagctgttccagaggaaccccaactacgaattgaaataatga 1080  
Dh 1446 agcttatgtcaaaagttagacagctgttccagaggaaccccaactacgaattgaaataatga 1505  
Qy 1081 agagagcggaagactatgatatactatctactgtatcttgaattgatagtgtgcagaattga 1140  
Dh 1506 agagagcggaagactatgatatactatctactgtatcttgaattgatagtgtgcagaattga 1565  
Qy 1141 tgatgcacaactcctcctccttaccacatctgcctcagttgcgtccagaagaagcattccaaac 1200  
Dh 1566 tgatgcacaactcctcctccttaccacatctgcctcagttgcgtccagaagaagcattccaaac 1625  
Qy 1201 ttgggtacatatacatctgcctgcgaagaagagagactggagactatgctcccttagtctcgc 1260  
Dh 1626 ttgggtacatatacatctgcctgcgaagaagagagactggagactatgctcccttagtctcgc 1685  
Qy 1261 ccccggtgcagagaagtataaagaatcaatttggaaacatggccctcaagcgatgtgtag 1320  
Dh 1686 ccccggtgcagagaagtataaagaatcaatttggaaacatggccctcaagcgatgtgtag 1745  
Qy 1321 gaagtaacaaaagctccgactttagcatacacagatgaaacctttagactctgtagac 1380  
Dh 1746 gaagtaacaaaagctccgactttagcatacacagatgaaacctttagactctgtagac 1805  
Qy 1381 tattcagcatgaaatcaggaactcttgggaaccttacttattggggaaagtltggagacacact 1440  
Dh 1806 tattcagcatgaaatcaggaactcttgggaaccttacttattggggaaagtltggagacacact 1865  
Qy 1441 gttgtattatattaaagaatcaagaacagacacataataatcaccctcaacggatcac 1500  
Dh 1866 gttgtattatattaaagaatcaagaacagacacataataatcaccctcaacggatcac 1925  
Qy 1501 tgatgcctgcctcttgcattcaagaagatataccaaaagtgtaaaaacttgaagaagatt 1560  
Dh 1926 tgatgcctgcctcttgcattcaagaagatataccaaaagtgtaaaaacttgaagaagatt 1985  
Qy 1561 tccaattctgcagagagaataattcaataataatggacagtgcataagaagtgggccc 1620  
Dh 1986 tccaattctgcagagagaataattcaataataatggacagtgcataagaagtgggccc 2045  
Qy 1621 aactaaatcagatccctgcgtgcgtgcagccgctatactactagttctgttaaatgagagag 1680  
Dh 2046 aactaaatcagatccctgcgtgcgtgcagccgctatactactagttctgttaaatgagagag 2105  
Qy 1681 agatctagcttcagagactcatctggccctctccatctgcatacagaagaatctgtatga 1740  
Dh 2106 agatctagcttcagagactcatctggccctctccatctgcatacagaagaatctgtatga 2165  
Qy 1741 aagaggaagaaccagataatgtccagacaagaagatgtcatccctgtttctgtatattgatga 1800  
Dh 2166 aagaggaagaaccagataatgtccagacaagaagatgtcatccctgtttctgtatattgatga 2225  
Qy 1801 gaacccaagaactgttacctccagagaataatatacaagcgtttctccccaatccaagcttgag 1860  
Dh 2226 gaacccaagaactgttacctccagagaataatatacaagcgtttctccccaatccaagcttgag 2285  
Qy 1861 ggaagcttgagatccagagatttccaagcctccaacatcatgacagacataatgtgctatgt 1920  
Dh 2286 ggaagcttgagatccagagatttccaagcctccaacatcatgacagacataatgtgctatgt 2345  
Qy 1921 ctttgatagtttgagattgtcagattgttgcataagagtgagcatatcgtatcatctaaag 1980  
Dh 2346 ctttgatagtttgagattgtcagattgttgcataagagtgagcatatcgtatcatctaaag 2405  
Qy 1981 catctgagacaagaactgactcctttctgtctcttctctcgatataacttcaaacacaa 2040  
Dh 2406 catctgagacaagaactgactcctttctgtctcttctctcgatataacttcaaacacaa 2465

Qy 2041 aatgtctatgaagaacacacttcaaccttaccattccattctcagagaagaactgttctcatg 2100  
Dh 2466 aatgtctatgaagaacacacttcaaccttaccattccattctcagagaagaactgttctcatg 2525  
Qy 2101 gatygaaaacccaggtcttatgattctgtgggtgcacaaactcgaacttctcggagaagag 2160  
Dh 2526 gatygaaaacccaggtcttatgattctgtgggtgcacaaactcgaacttctcggagaagag 2585  
Qy 2161 catgaccgcttacttgaaagtttctagtttgtaacaagaacactggtgattattacgaaga 2220  
Dh 2586 catgaccgcttacttgaaagtttctagtttgtaacaagaacactggtgattattacgaaga 2645  
Qy 2221 cagttatgaagaatttcaacatacttgcctgagtaaaaaacaatgcatgtgaaccaagaag 2280  
Dh 2646 cagttatgaagaatttcaacatacttgcctgagtaaaaaacaatgcatgtgaaccaagaag 2705  
Qy 2281 ctctccagaagaattcaagacacctgagcactaaggcaaaagcaatttaatgcccaccac 2340  
Dh 2706 ctctccagaagaattcaagacacctgagcactaaggcaaaagcaatttaatgcccaccac 2709  
Qy 2341 agtctgaaagcgcatcaacggaataactcgtactactcttcagttcagatcaagaaga 2400  
Dh 2710 agtctgaaagcgcatcaacggaataactcgtactactcttcagttcagatcaagaaga 2747  
Qy 2401 aattgcatatgatgatataccatatacagttgaaatgaaagaagaagatttgcataattatga 2460  
Dh 2748 aattgcatatgatgatataccatatacagttgaaatgaaagaagaagatttgcataattatga 2807  
Qy 2461 tgagatgaaataatcagaagcccccgagcttcaaaaagaacaacgacactattttatg 2520  
Dh 2808 tgagatgaaataatcagaagcccccgagcttcaaaaagaacaacgacactattttatg 2867  
Qy 2521 tgacagtgagaaggtctcgggattatagatgagtaagcttcccaacatggttcaagaagag 2580  
Dh 2868 tgacagtgagaaggtctcgggattatagatgagtaagcttcccaacatggttcaagaagag 2927  
Qy 2581 ggtcagagtgagcagtgctccctcagttcaagaagaagtgttttccagaagaatttactgag 2640  
Dh 2928 ggtcagagtgagcagtgctccctcagttcaagaagaagtgttttccagaagaatttactgag 2987  
Qy 2641 ctcccttaccagcccttataccgttggagaactaaatgaaacatttgggctccctgggccc 2700  
Dh 2988 ctcccttaccagcccttataccgttggagaactaaatgaaacatttgggctccctgggccc 3047  
Qy 2701 atataagaagcagaagttgaagaataatcatgatacttcaagaatcaacagccctcgc 2760  
Dh 3048 atataagaagcagaagttgaagaataatcatgatacttcaagaatcaacagccctcgc 3107  
Qy 2761 tccctattccttatactctagccttattcttatgagaagaatcagaagcagaagcaga 2820  
Dh 3108 tccctattccttatactctagccttattcttatgagaagaatcagaagcagaagcaga 3167  
Qy 2821 aactagaanaaactttgtcaagccttaatgaaacaaaacttacttttgaagaatgaaaga 2880  
Dh 3168 aactagaanaaactttgtcaagccttaatgaaacaaaacttacttttgaagaatgaaaga 3227  
Qy 2881 tcaatagcccccacataaagatgatttgactgcaagaagccttggtctatttctctatgt 2940  
Dh 3228 tcaatagcccccacataaagatgatttgactgcaagaagccttggtctatttctctatgt 3287  
Qy 2941 tgacctgaaaaaagatgttgactcaagccttgatgtgaacccctctggtctgcacacata 3000  
Dh 3288 tgacctgaaaaaagatgttgactcaagccttgatgtgaacccctctggtctgcacacata 3347  
Qy 3001 cacactgaaacccctgctcatatggagagaacagtgcacagtacagaatttgcctgtttttac 3060  
Dh 3348 cacactgaaacccctgctcatatggagagaacagtgcacagtacagaatttgcctgtttttac 3407  
Qy 3061 catctttgatgagacaaagaagcgtgatactcactgtaaaatataatgaaagaagaactgcagagc 3120  
Dh 3408 catctttgatgagacaaagaagcgtgatactcactgtaaaatataatgaaagaagaactgcagagc 3467



QY 3121 tccctgcaatataccaga tggagaagatcccaacttttaagaagaatatacgtctccatgcaat 3180  
|||||  
Db 3468 tccctgcaatataccaga tggagaagatcccaacttttaagaagaatatacgtctccatgcaat 3527  
QY 3181 caatgctacaataatgataatacaatacctgcttagtaaatggtccaga tcaaaaggatcgc 3240  
|||||  
Db 3528 caatgctacaataatgataatacaatacctgcttagtaaatggtccaga tcaaaaggatcgc 3587  
QY 3241 atggtatctgctcaagcattggcagcaatgaaacatccattctatatttccatggtgaca 3300  
|||||  
Db 3588 atggtatctgctcaagcattggcagcaatgaaacatccattctatatttccatggtgaca 3647  
QY 3301 tgtgttcaactgctacgaaataaagggaggtataaatggtcactgtacatctctacacag 3360  
|||||  
Db 3648 tgtgttcaactgctacgaaataaagggaggtataaatggtcactgtacatctctacacag 3707  
QY 3361 tgtttttgagacagtggaatgttaacatccaaagctggaattggcgggtggaatgct 3420  
|||||  
Db 3708 tgtttttgagacagtggaatgttaacatccaaagctggaattggcgggtggaatgct 3767  
QY 3421 tattggcgagacattacattgcttgatgagacacatttttctgtgtgacagcaataatg 3480  
|||||  
Db 3768 tattggcgagacattacattgcttgatgagacacatttttctgtgtgacagcaataatg 3827  
QY 3481 tcaagctcccttggaatggtctctgacacatagagatttccaga ttaacagcttcag 3540  
|||||  
Db 3828 tcaagctcccttggaatggtctctgacacatagagatttccaga ttaacagcttcag 3887  
QY 3541 acaatatgagcagttgggcccnaaagctggcagacattcattatccgcatcaatcaatgc 3600  
|||||  
Db 3888 acaatatgagcagttgggcccnaaagctggcagacattcattatccgcatcaatcaatgc 3947  
QY 3601 ctggagacacaaagagccctttcttgatcagaatggaatctgttgacacaaatgtatat 3660  
|||||  
Db 3948 ctggagacacaaagagccctttcttgatcagaatggaatctgttgacacaaatgtatat 4007  
QY 3661 tcaagcagatcaaaagccaggtggccgtcagaaagttctccagcctctacatctccatgc 3720  
|||||  
Db 4008 tcaagcagatcaaaagccaggtggccgtcagaaagttctccagcctctacatctccatgc 4067  
QY 3721 tatcatcatgtatagcttgaatgggaagaagtggcagacattatccgaaatccactgg 3780  
|||||  
Db 4068 tatcatcatgtatagcttgaatgggaagaagtggcagacattatccgaaatccactgg 4127  
QY 3781 aaacttaatggtcttcttcttgcaatgtggaattatctggaataaaacaatatlttttaa 3840  
|||||  
Db 4128 aaacttaatggtcttcttcttgcaatgtggaattatctggaataaaacaatatlttttaa 4187  
QY 3841 cccctcaatatgtctcgatatacctcgtttgcaacccaacatcatatagcatctgcagcac 3900  
|||||  
Db 4188 cccctcaatatgtctcgatatacctcgtttgcaacccaacatcatatagcatctgcagcac 4247  
QY 3901 tcttcgacatgagatgtaatgggtctgtatataaagtgtcagcatgccaattgggaatgga 3960  
|||||  
Db 4248 tcttcgacatgagatgtaatgggtctgtatataaagtgtcagcatgccaattgggaatgga 4307  
QY 3961 gaaataagaataatcagatcacagatattgattatccctacttaacaaatgtttgc 4020  
|||||  
Db 4308 gaaataagaataatcagatcacagatattgattatccctacttaacaaatgtttgc 4367  
QY 4021 caactggtctcttcaaaaagctcgacttcaaccccaagggaggaatlaatccctggagacc 4080  
|||||  
Db 4368 caactggtctcttcaaaaagctcgacttcaaccccaagggaggaatlaatccctggagacc 4427  
QY 4081 tcaatggtgataatcccaaaagatggtctgcaagtggacttccagaagaacaaatgaagtcaac 4140  
|||||  
Db 4428 tcaatggtgataatcccaaaagatggtctgcaagtggacttccagaagaacaaatgaagtcaac 4487  
QY 4141 aggaatgaactaccaggaatlaaataatctgcttaccagatgtatgtgagaagtctcct 4200  
|||||  
Db 4488 aggaatgaactaccaggaatlaaataatctgcttaccagatgtatgtgagaagtctcct 4547  
QY 4201 catctccagcagtcagatggtccatcagtgactctcttcttcagaatggtcaaaatlaaa 4260

Db 4548 catctccagcagtcagatggtccatcagtgactctcttcttcagaatggtcaaaatlaaa 4607  
|||||  
QY 4261 ggttttcagggagaatcaagaatcctctcaacatctgtgtgaactctctagaccacgcgt 4320  
|||||  
Db 4608 ggttttcagggagaatcaagaatcctctcaacatctgtgtgaactctctagaccacgcgt 4667  
QY 4321 actgaactgcctactcttccgaatcccccagagttgggtgcacaaatggtccctggagat 4380  
|||||  
Db 4668 actgaactgcctactcttccgaatcccccagagttgggtgcacaaatggtccctggagat 4727  
QY 4381 ggaagtctgtggtgcgagggcacaggaacctctactga 4417  
|||||  
Db 4728 ggaagtctgtggtgcgagggcacaggaacctctactga 4764

RESULT 7  
US-09-364-862-13  
; Sequence 137 Application US/09364862  
; Patent No. 6221349  
; GENERAL INFORMATION:  
; APPLICANT: Coulo, Linda B.  
; APPLICANT: Colosi, Peter C.  
; TITLE OF INVENTION: ADEMO-ASSOCIATED VECTORS FOR EXPRESSION OF FACTOR VIII  
; TITLE OF INVENTION: BY TARGET  
; TITLE OF INVENTION: CELLS  
; FILE REFERENCE: AVIGEN-03743  
; CURRENT APPLICATION NUMBER: US/09/364, 862  
; EARLIER FILING DATE: 1999-07-30  
; EARLIER APPLICATION NUMBER: 60/125, 974  
; EARLIER FILING DATE: 1999-03-24  
; EARLIER APPLICATION NUMBER: 60/104, 994  
; EARLIER FILING DATE: 1998-10-20  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: Patentln Ver. 2.0  
; SEQ ID NO 13  
; LENGTH: 11933  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
US-09-364-862-13

Query Match 96.2%; Score 4249.4; DB 4; Length 11933;  
Best Local Similarity 98.2%; Pred. No. 0;  
Matches 4338; Conservative 0; Mismatches 1; Indels 78; Gaps 1;

QY 1 catgcaaatagagctccacacgtcttctctgtgctcttgagattcgattctgattatgctc 60  
|||||  
Db 426 catgcaaatagagctccacacgtcttctctgtgctcttgagattcgattctgattatgctc 485  
QY 61 caccagaagaatactactcctgggtgagatggaactgtcatggaactatataatgcaaatgact 120  
|||||  
Db 486 caccagaagaatactactcctgggtgagatggaactgtcatggaactatataatgcaaatgact 545  
QY 121 cgggtggtgctcctgtgtgagcagaagatttctcctctaagtgagcacaatctttccattcaa 180  
|||||  
Db 546 cgggtggtgctcctgtgtgagcagaagatttctcctctaagtgagcacaatctttccattcaa 605  
QY 181 caactaagtcgtgtacaaaagacactgtttgtgagaatcaacggttcaacttccaacat 240  
|||||  
Db 606 caactaagtcgtgtgtacaaaagacactgtttgtgagaatcaacggttcaacttccaacat 665  
QY 241 cgtctaaagccaagccacccctgagttggtctgctaggtctctacacatcccaagctgaagttta 300  
|||||  
Db 666 cgtctaaagccaagccacccctgagttggtctgctaggtctctacacatcccaagctgaagttta 725  
QY 301 tgataagtggtgtataacttaagaacatggttcccatcctgctgaagcttcttatgtctgt 360  
|||||  
Db 726 tgataagtggtgtataacttaagaacatggttcccatcctgctgaagcttcttatgtctgt 785  
QY 361 tgggtatcctactgtgaagcttctgagggagctgataatgatgatcagacagtcagtaaaag 420



```

Db      786  tggctgctactcgtgaagctctcgtgaggagctggaatgatgatcgcgaacgtcaag  845
OY      421  ggggaaagaagatgataaagctctccctgggtggaagccatcatcatgtctgagctct  480
Db      846  ggggaaagaagatgataaagctctccctgggtggaagccatcatcatgtctgagctct  905
OY      481  gaaagaatcgtctcaatcgtgctctgacccactgctgaccttaactcatatctcttca  540
Db      906  gaaagagaatcgtctcaatcgtgctctgacccactgctgaccttaactcatatctcttca  965
OY      541  tctggagactcgttaaaagacttgaatcagagctcatttggagccctactcaagtatagga  600
Db      966  tctggagactcgttaaaagacttgaatcagagctcatttggagccctactcaagtatagga  1025
OY      601  agggagctctggccaaaggaaaagacacagagaccttgacaaaatttaactcttctgctg  660
Db      1026  agggagctctggccaaaggaaaagacacagagaccttgacaaaatttaactcttctgctg  1085
OY      661  attctgataagggaagaaagtctggcactcagaagaaagaaaccccttgatgacagatagga  720
Db      1086  attctgataagggaagaaagtctggcactcagaagaaagaaaccccttgatgacagatagga  1145
OY      721  tgcctgcactcgcctcgggctcggcttaaaatgacacagtcataatggttatgttaacaggtc  780
Db      1146  tgcctgcactcgcctcgggctcggcttaaaatgacacagtcataatggttatgttaacaggtc  1205
OY      781  tctggcagagctgatactgatagcccacagaagaaatcagctatgagcatgatactggaatgg  840
Db      1206  tctggcagagctgatactgatagcccacagaagaaatcagctatgagcatgatactggaatgg  1265
OY      841  caccactcctgaagtgactcactcaatatctcctcgaaggtcacacattctctgtgagaacca  900
Db      1266  caccactcctgaagtgactcactcaatatctcctcgaaggtcacacattctctgtgagaacca  1325
OY      901  tggccagaggtctccttggaaatctcgcgaataacttctctctgctcaaaaccttgat  960
Db      1326  tggccagaggtctccttggaaatctcgcgaataacttctctctgctcaaaaccttgat  1385
OY      961  gggcctctgacagcttctactgcttctgcatatctctcctcacaacaatgatactgcatgga  1020
Db      1386  gggcctctgacagcttctactgcttctgcatatctctcctcacaacaatgatactgcatgga  1445
OY      1021  agcttatgtccaagtagacagctgttccagaggaaccccaactacgaatgaaataatgaa  1080
Db      1446  agcttatgtccaagtagacagctgttccagaggaaccccaactacgaatgaaataatgaa  1505
OY      1081  aggaagcgggaagactatgatagtatcttactgattcttggaaatgatagtgtgcaagtttga  1140
Db      1506  aggaagcgggaagactatgatagtatcttactgattcttggaaatgatagtgtgcaagtttga  1565
OY      1141  tgaatgaacactctcctccttcaatccaaatcgtcagtttgccaagaagcatctctaaac  1200
Db      1566  tgaatgaacactctcctccttcaatccaaatcgtcagtttgccaagaagcatctctaaac  1625
OY      1201  ctgggtacatcatcattgctgctcgtgaagagagagactgggactatgctccctagtctctgc  1260
Db      1626  ctgggtacatcatcattgctgctcgtgaagagagagactgggactatgctccctagtctctgc  1685
OY      1261  ccccgatgacagaagtataaagtcaattttgaacaatgagccctcagcgattgtgtag  1320
Db      1686  ccccgatgacagaagtataaagtcaattttgaacaatgagccctcagcgattgtgtag  1745
OY      1321  gaagtaacaaaagctccgacttlatgtgcatcacagatgaaacctttaagactcgtgaagc  1380
Db      1746  gaagtaacaaaagctccgacttlatgtgcatcacagatgaaacctttaagactcgtgaagc  1805
OY      1381  taatcagcatgaaatcaggaatcttgggaaccttaacttatactgaggaggttggagacacact  1440
Db      1806  taatcagcatgaaatcaggaatcttgggaaccttaacttatactgaggaggttggagacacact  1865
OY      1441  gtgtattatatattaaagatcaagaagcagaccataaactctaacctcaggaattcac  1500
Db      1866  gtgtattatatattaaagatcaagaagcagaccataaactctaacctcaggaattcac  1925
OY      1501  tgaatccgtctccttctgatactcaagagatttaaccaaaaggtgtaaacatttgaagattt  1560
Db      1926  tgaatccgtctccttctgatactcaagagatttaaccaaaaggtgtaaacatttgaagattt  1985
OY      1561  tccaattctgcagagagaataltcaaatataaataatgacagctgactgtagaagatgggcc  1620
Db      1986  tccaattctgcagagagaataltcaaatataaataatgacagctgactgtagaagatgggcc  2045
OY      1621  aactaaatcagatctcgtggtgctgagccgtattactctagtttctgttaatttgaagag  1680
Db      2046  aactaaatcagatctcgtggtgctgagccgtattactctagtttctgttaatttgaagag  2105
OY      1681  agatcagcttcagagactcattggtccctccctcatatcgtgtataaagaatctataatca  1740
Db      2106  agatcagcttcagagactcattggtccctccctcatatcgtgtataaagaatctataatca  2165
OY      1741  aagaagaaaccagataatcgtcagaacaagagaaatcgtcatccgttctgtatattgata  1800
Db      2166  aagaagaaaccagataatcgtcagaacaagagaaatcgtcatccgttctgtatattgata  2225
OY      1801  gaaccgaagctggtacactccaagagaaatatacaagcttctccccaatccagctgagat  1860
Db      2226  gaaccgaagctggtacactccaagagaaatatacaagcttctccccaatccagctgagat  2285
OY      1861  gcaacttgagatccagaagttccaaagctcccaatcatatgacacagcatcaatgtgcatgt  1920
Db      2286  gcaacttgagatccagaagttccaaagctcccaatcatatgacacagcatcaatgtgcatgt  2245
OY      1921  ttttgataagtttgcaatcgtcagttgttctgtgcatagaggtggcactggtatcttaag  1980
Db      2346  ttttgataagtttgcaatcgtcagttgttctgtgcatagaggtggcactggtatcttaag  2405
OY      1981  cattggagacagagactgaactctcttctgctctcttctcctggatataaccttcaaacaca  2040
Db      2406  cattggagacagagactgaactctcttctgctctcttctcctggatataaccttcaaacaca  2465
OY      2041  aatgctcattgaagaacacactcaacctatactccattctccagaggaactgtctctcatgtc  2100
Db      2466  aatgctcattgaagaacacactcaacctatactccattctccagaggaactgtctctcatgtc  2525
OY      2101  gatvgaaaaccaggtctatagatctctggggtgcgaacactcagactctcggaacagagg  2160
Db      2526  gatvgaaaaccaggtctatagatctctggggtgcgaacactcagactctcggaacagagg  2585
OY      2161  catgacgccttactggaagtttctagtttctgtaacaagaacacgtggtattatcagaga  2220
Db      2586  catgacgccttactggaagtttctagtttctgtaacaagaacacacgtggtattatcagaga  2645
OY      2221  cagttatgaagataattcaagcatactgtcgtgagtaaaacaatgcatcttgaaaccaagag  2280
Db      2646  cagttatgaagataattcaagcatactgtcgtgagtaaaacaatgcatcttgaaaccaagag  2705
OY      2281  ctctctccagaattcaagaacacctagcactagggcaaaagcaatttaatgcaaccccaac  2340
Db      2706  ctctctccagaattcaagaacacctagcactagggcaaaagcaatttaatgcaaccccaac  2709
OY      2341  agcttgaaagccatcaacgaggaataactcgttaactcactcctcagttcagatcaagaga  2400
Db      2710  agcttgaaagccatcaacgaggaataactcgttaactcactcctcagttcagatcaagaga  2747
OY      2401  aattgacatagatgataccataltcagttgaaatgaaagaaagattttgacattatga  2460
Db      2748  aattgacatagatgataccataltcagttgaaatgaaagaaagattttgacattatga  2807
OY      2461  tgaagatgaaatcaagagcccgagcttcaaaagaaacacagacactatttatgtc  2520
Db      2808  tgaagatgaaatcaagagcccgagcttcaaaagaaacacagacactatttatgtc  2867
OY      2521  tgcagtggaagagctctgggataatgagatgagatgagctcccaatggtttcctaaagag  2580
Db      2868  tgcagtggaagagctctgggataatgagatgagatgagctcccaatggtttcctaaagag  2927
```



QY 2581 ggcacagatggcagctgcccctcagttcagaagaagttgtttccagaattactatg 2640  
|||||  
Db 2928 ggcctcagagtgagcagtgccctcagttcagaagaagttgtttccagaattactatg 2887  
QY 2641 cctccttactcagcccttataccgtggagaactaaatgaacatttggactcctgggccc 2700  
|||||  
Db 2988 cctccttactcagcccttataccgtggagaactaaatgaacatttggactcctgggccc 3047  
QY 2701 atatatagaagcagaagtttgaagaataatcatcgttaactttcagaatacaggcctctcg 2760  
|||||  
Db 3048 atataagaagcagaagtttgaagaataatcatcgttaactttcagaatacaggcctctcg 3107  
QY 2761 tccctatcctcttactcagccttatttcttcttgaaggaatcagaaggagaagcaga 2820  
|||||  
Db 3108 tccctatcctcttactcagccttatttcttcttgaaggaatcagaaggagaagcaga 3167  
QY 2821 acctagaanaaaacttctcagccttaatgaacccaaacttacttcttgaagaatgcaaca 2880  
|||||  
Db 3168 acctagaanaaaacttctcagccttaatgaacccaaacttacttcttgaagaatgcaaca 3227  
QY 2881 tcatatggcccccacccaagaatgagtttgaatcagaagcctgggttatttctctatgt 2940  
|||||  
Db 3228 tcatatggcccccacccaagaatgagtttgaatcagaagcctgggttatttctctatgt 3287  
QY 2941 tgaacctggaanaaagatgtgacacccaagcctggaatcagccctctggtctgcacacaa 3000  
|||||  
Db 3288 tgaacctggaanaaagatgtgacacccaagcctggaatcagccctctggtctgcacacaa 3347  
QY 3001 cacactgaacccctgtcatalgagagacaaatgacagtaacaggaatttgcctgtttttcac 3060  
|||||  
Db 3348 cacactgaacccctgtcatalgagagacaaatgacagtaacaggaatttgcctgtttttcac 3407  
QY 3061 catcttctgtagagacccaagaagcgtgacttccactgtaataatgagaagaactgcagggc 3120  
|||||  
Db 3408 catcttctgtagagacccaagaagcgtgacttccactgtaataatgagaagaactgcagggc 3467  
QY 3121 tccctcgaatccagatggaagaatcccaactttaaagaagaattacgcgtccatcagat 3180  
|||||  
Db 3468 tccctcgaatccagatggaagaatcccaactttaaagaagaattacgcgtccatcagat 3527  
QY 3181 caatggctacataatgatacactaccctggttagtaatgagtcacagatcaaaagattcg 3240  
|||||  
Db 3528 caatggctacataatgatacactaccctggttagtaatgagtcacagatcaaaagattcg 3587  
QY 3241 atggtatcgtctcagcagatggcgacgaatgaaacatccatcttcatcttcaagtggaca 3300  
|||||  
Db 3588 atggtatcgtctcagcagatggcgacgaatgaaacatccatcttcatcttcaagtggaca 3647  
QY 3301 tgtgttcactgtagcaaaaaaagagagataaaatgagcctgttagaatctctacag 3360  
|||||  
Db 3648 tgtgttcactgtagcaaaaaaagagagataaaatgagcctgttagaatctctacag 3707  
QY 3361 tgtttttgacagtggaatgtttacatcccaagcttgaatttggcgggttgaatgctc 3420  
|||||  
Db 3708 tgtttttgacagtggaatgtttacatcccaagcttgaatttggcgggttgaatgctc 3767  
QY 3421 tatgtggcagcattcatcagcttgggagtggagacactttttctgtgttacaagcaataatg 3480  
|||||  
Db 3768 tatgtggcagcattcatcagcttgggagtggagacactttttctgtgttacaagcaataatg 3827  
QY 3481 tcagactcccttgggaatggtctctgacacattagaattttcagaattacaactcag 3540  
|||||  
Db 3828 tcagactcccttgggaatggtctctgacacattagaattttcagaattacaactcag 3587  
QY 3541 acaatatgacagtgggccccaagaagctggccagacttcatattccggatcaatcaatgc 3600  
|||||  
Db 3888 acaatatgacagtgggccccaagaagctggccagacttcatattccggatcaatcaatgc 3947  
QY 3601 ctggagacccaagaagccctttctctgtagtcaagtgatcgttggcaccaatgattat 3660  
|||||  
Db 3948 ctggagacccaagaagccctttctctgtagtcaagtgatcgttggcaccaatgattat 4007

QY 3661 tcacggcatcaagaccacaaggtgcccctcagaagttctccagcctctacatctcagtt 3720  
|||||  
Db 4008 tcacggcatcaagaccacaaggtgcccctcagaagttctccagcctctacatctcagtt 4067  
QY 3721 tatcatcatgtagtcttcttgaatggagaagtgagcagacttaccgagaattccactg 3780  
|||||  
Db 4068 tatcatcatgtagtcttcttgaatggagaagtgagcagacttaccgagaattccactg 4127  
QY 3781 aaacttaatggtcttcttcttggcgaatgtagtattcatcttggaataaacacataatttta 3840  
|||||  
Db 4128 aaacttaatggtcttcttcttggcgaatgtagtattcatcttggaataaacacataatttta 4187  
QY 3841 cccctcaatattgtcgcgatacatccggttggacccaacacttatagacttgcagcac 3900  
|||||  
Db 4188 cccctcaatattgtcgcgatacatccggttggacccaacacttatagacttgcagcac 4247  
QY 3901 tcttcgcatgagttgagtggtcgtgtattaaatgltgacatgccaatgggaattgga 3960  
|||||  
Db 4248 tcttcgcatgagttgagtggtcgtgtattaaatgltgacatgccaatgggaattgga 4307  
QY 3961 gagtaagaacataatcagatgacagattactgttcatcttacttaccataatgtttgc 4020  
|||||  
Db 4308 gagtaagaacataatcagatgacagattactgttcatcttacttaccataatgtttgc 4367  
QY 4021 cacctgtctccttcaaaaagctcgacttcaacctccaaggagagatgaatgcctggagacc 4080  
|||||  
Db 4368 cacctgtctccttcaaaaagctcgacttcaacctccaaggagagatgaatgcctggagacc 4427  
QY 4081 tcagtgtaaatccaaagaatggtgctgcaagtggacttccagaagaacatgaagaatcac 4140  
|||||  
Db 4428 tcagtgtaaatccaaagaatggtgctgcaagtggacttccagaagaacatgaagaatcac 4487  
QY 4141 aggaagtaactactcagaaggttaaaatctcgtctaccagaatgtagtgaagaagttcct 4200  
|||||  
Db 4488 aggaagtaactactcagaaggttaaaatctcgtctaccagaatgtagtgaagaagttcct 4547  
QY 4201 catctccagcagtcagaatgagccatgagtcgactcctttttcagaatggcaagtaaa 4260  
|||||  
Db 4548 catctccagcagtcagaatgagccatgagtcgactcctttttcagaatggcaagtaaa 4607  
QY 4261 ggttttcagggaatacgaactccttccacactgtgtgtaactctcttagaccacggt 4320  
|||||  
Db 4608 ggttttcagggaatacgaactccttccacactgtgtgtaactctcttagaccacggt 4667  
QY 4321 actgactcgtactcctcgaaattcaaccccaagatgtgtgacacagatggcctgaagat 4380  
|||||  
Db 4668 actgactcgtactcctcgaaattcaaccccaagatgtgtgacacagatggcctgaagat 4727  
QY 4381 ggaagttcggggtcgagagcagagacccttactga 4417  
|||||  
Db 4728 ggaagttcggggtcgagagcagagacccttactga 4764

RESULT 8  
US-08-882-083-1  
: Sequence 1, Application US/08882083  
: Patent No. 5869292  
: GENERAL INFORMATION:  
: APPLICANT: VOORBERG, Johannes J.  
: TITLE OF INVENTION: HYBRID PROTEINS WITH MODIFIED ACTIVITY  
: NUMBER OF SEQUENCES: 17  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: Foley & Lardner  
: STREET: 3000 K Street, N.W., Suite 500  
: CITY: Washington  
: STATE: D.C.  
: COUNTRY: USA  
: ZIP: 20007-5109  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: PatentIn Release #1.0, Version #1.30



```

: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/882,083
: FILING DATE:
: CLASSIFICATION: 514
: PRIORITY APPLICATION DATA:
: APPLICATION NUMBER: US 08/558,107
: FILING DATE: 13-NOV-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: ISACSON, John P.
: REGISTRATION NUMBER: 33,715
: REFERENCE/DOCKET NUMBER: 30472/212
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (202)672-5300
: TELEFAX: (202)672-5399
: TELEX: 904136
: INFORMATION FOR SEQ. ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 5035 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 35..5017
: US-08-882-083-1

```

Query Match 84.8% Score 3743.8 DB 2 Length 5035:

Best Local Similarity 87.5% Pred. No. 0; Mismatches 52; Indels 570; Gaps 2;

Matches 4355; Conservative 0; Mismatches 52; Indels 570; Gaps 2;

```

QY 1 catgcaataagaactcctcaccctgctctctctgctgctcttctgcaatctgcttaagtc 60
DB 34 catggaatagactctccacactgctctctctgctgctcttctgcaatctgcttaagtc 93
QY 61 caccagaagatactactctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 120
DB 94 caccagaagatactactctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 153
QY 121 cggctgagctcctgctgagcgaagatctcctcctgagctgcaaatcttccatcaa 180
DB 154 cggctgagctcctgctgagcgaagatctcctcctgagctgcaaatcttccatcaa 213
QY 181 caccctagctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 240
DB 214 caccctagctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 273
QY 241 cggctgagctcctgctgagcgaagatctcctcctgagctgcaaatcttccatcaa 300
DB 274 cggctgagctcctgctgagcgaagatctcctcctgagctgcaaatcttccatcaa 333
QY 301 tgaataagctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 360
DB 334 tgaataagctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 393
QY 361 tggctgagctcctgctgagcgaagatctcctcctgagctgcaaatcttccatcaa 420
DB 394 tggctgagctcctgctgagcgaagatctcctcctgagctgcaaatcttccatcaa 453
QY 421 tggctgagctcctgctgagcgaagatctcctcctgagctgcaaatcttccatcaa 480
DB 454 tggctgagctcctgctgagcgaagatctcctcctgagctgcaaatcttccatcaa 513
QY 481 tggctgagctcctgctgagcgaagatctcctcctgagctgcaaatcttccatcaa 540
DB 514 tggctgagctcctgctgagcgaagatctcctcctgagctgcaaatcttccatcaa 573
QY 541 tggctgagctcctgctgagcgaagatctcctcctgagctgcaaatcttccatcaa 600
DB 574 tggctgagctcctgctgagcgaagatctcctcctgagctgcaaatcttccatcaa 633
QY 601 tggctgagctcctgctgagcgaagatctcctcctgagctgcaaatcttccatcaa 660

```

```

DB 634 agggagctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 693
QY 661 attgataagaggaagattgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 720
DB 694 attgataagaggaagattgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 753
QY 721 tggctgagctcctgctgagcgaagatctcctcctgagctgcaaatcttccatcaa 780
DB 754 tggctgagctcctgctgagcgaagatctcctcctgagctgcaaatcttccatcaa 813
QY 781 tggctgagctcctgctgagcgaagatctcctcctgagctgcaaatcttccatcaa 840
DB 814 tggctgagctcctgctgagcgaagatctcctcctgagctgcaaatcttccatcaa 873
QY 841 caccctagctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 900
DB 874 caccctagctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 933
QY 901 tggctgagctcctgctgagcgaagatctcctcctgagctgcaaatcttccatcaa 960
DB 934 tggctgagctcctgctgagcgaagatctcctcctgagctgcaaatcttccatcaa 993
QY 961 tggctgagctcctgctgagcgaagatctcctcctgagctgcaaatcttccatcaa 1020
DB 994 tggctgagctcctgctgagcgaagatctcctcctgagctgcaaatcttccatcaa 1053
QY 1021 tggctgagctcctgctgagcgaagatctcctcctgagctgcaaatcttccatcaa 1080
DB 1054 tggctgagctcctgctgagcgaagatctcctcctgagctgcaaatcttccatcaa 1113
QY 1081 tggctgagctcctgctgagcgaagatctcctcctgagctgcaaatcttccatcaa 1140
DB 1114 tggctgagctcctgctgagcgaagatctcctcctgagctgcaaatcttccatcaa 1173
QY 1141 tggctgagctcctgctgagcgaagatctcctcctgagctgcaaatcttccatcaa 1200
DB 1174 tggctgagctcctgctgagcgaagatctcctcctgagctgcaaatcttccatcaa 1233
QY 1201 tggctgagctcctgctgagcgaagatctcctcctgagctgcaaatcttccatcaa 1260
DB 1234 tggctgagctcctgctgagcgaagatctcctcctgagctgcaaatcttccatcaa 1293
QY 1261 tggctgagctcctgctgagcgaagatctcctcctgagctgcaaatcttccatcaa 1320
DB 1294 tggctgagctcctgctgagcgaagatctcctcctgagctgcaaatcttccatcaa 1353
QY 1321 tggctgagctcctgctgagcgaagatctcctcctgagctgcaaatcttccatcaa 1380
DB 1354 tggctgagctcctgctgagcgaagatctcctcctgagctgcaaatcttccatcaa 1413
QY 1381 tggctgagctcctgctgagcgaagatctcctcctgagctgcaaatcttccatcaa 1440
DB 1414 tggctgagctcctgctgagcgaagatctcctcctgagctgcaaatcttccatcaa 1473
QY 1441 tggctgagctcctgctgagcgaagatctcctcctgagctgcaaatcttccatcaa 1500
DB 1474 tggctgagctcctgctgagcgaagatctcctcctgagctgcaaatcttccatcaa 1533
QY 1501 tggctgagctcctgctgagcgaagatctcctcctgagctgcaaatcttccatcaa 1560
DB 1534 tggctgagctcctgctgagcgaagatctcctcctgagctgcaaatcttccatcaa 1593
QY 1561 tggctgagctcctgctgagcgaagatctcctcctgagctgcaaatcttccatcaa 1620
DB 1594 tggctgagctcctgctgagcgaagatctcctcctgagctgcaaatcttccatcaa 1653
QY 1621 tggctgagctcctgctgagcgaagatctcctcctgagctgcaaatcttccatcaa 1680
DB 1654 tggctgagctcctgctgagcgaagatctcctcctgagctgcaaatcttccatcaa 1713
QY 1681 tggctgagctcctgctgagcgaagatctcctcctgagctgcaaatcttccatcaa 1740
DB 1714 tggctgagctcctgctgagcgaagatctcctcctgagctgcaaatcttccatcaa 1773

```



1741 aagaggaaacccaataatgtcaagacaagagaaatgcatcctgttttctgtattga 1800  
1774 AAGAGAAACCAATATATGTCAAGACAAGAAATGATCTGTTTCTGTATTATGA 1833  
1801 gaaccgaactggttacctacagagaataatacaacgcttctcccaatccagctga 1860  
1834 GAACCAAGCTGTACTCTCAAGAGAATATACAAAGCTTCTCCCAATCCAGCTGAGT 1893  
1861 gcaagcttgagatccagagatcccaagcctcccaatcatatgcacagcataatg 1920  
1894 GCAAGCTTGAGATCCAGAGATCCCAAGCTTCCAAATCATATGACAGATCAATG 1953  
1921 ttctgaatgttgcagttctcagttgttctgaatgagtgagcactgttacctaa 1980  
1954 TTTTGAATGTTTGCATGTTGCTGCTTTGTTGCAATGAGTGGCACTGTGACTT 2013  
1981 catlgaagacagactgaactcctctctgtctctctctctgaatacacttcaaca 2040  
2014 CATGGAGCACAGACTGACTTCTCTCTCTCTCTCTCTGATATACCTTCAACACA 2073  
2041 aatgtctatgaagacaactcaacctatcccatctcaagagaactgcttcatgtc 2100  
2074 AATGCTATGAAGACACACTCACTATTCCTATCTCAGAGAAAGTCTTATGTC 2133  
2101 gatggaacccaagctctatgtatcttggtggtgccaaactcaagacttccgaaga 2160  
2134 GATGGAACCAACCAAGCTCTATGATCTGGGTGCCCAACTCAGACTTTCGGAACG 2193  
2161 catgaccgcttactgaaggttctcagttgttgaacaaactggtgatattacaaga 2220  
2194 CATGACCGCTTACTGAAAGCTTCTGATGTTGATTCAGAGGGGAGAGAGACAC 2253  
2221 cagttatgaagatattca-----gcatactgtcgtgaataaacaatgc 2265  
2254 TCTGACCTGAGAAAGATATTCAGTGAAGACGACGACTCATGACATGCTGACAG 2313  
2266 catlgaaccaaagaacttcccaagaatcccaagacccctagcactgagcaaaaga 2325  
2314 GATGAAACCAAGAGCTTCTCCAGATTCAGACACCTTAGCAGTAAAGCAAAAT 2373  
2326 taatgcac----- 2334  
2374 TATGACACACAAATTCAGAAATATGACATAGAGAGACTGACCTTGTGTCACAC 2433  
2335 ----- 2334  
2434 AACACCTATGCTAAATATCAAAATGTCTCTAGTATTTGTTGATGCTCTTGCA 2493  
2335 ----- 2334  
2494 GAGTCTACTCCACATGGGCTATCCTTATCTGATCTCCAGAAAGCCAAATAT 2553  
2335 ----- 2334  
2554 TTCTGATGATCATGACCTGAGCAATATAGAGTATATACAGCCTGTCTGAATG 2613  
2335 ----- 2334  
2614 CTTACAGGCACAGCTTCATCAGTGGGACATGATTTACCCCTGAGTACAGCT 2673  
2335 ----- 2334  
2674 AATTAAGATTAAATGAGAAACTGGGACAACCTGAGATCTCTTGTGGATTAAC 2733  
2335 ----- 2334  
2734 TGGTACTCAGATACCAAGAGAGTGAATCCCAAGAGAGTCCAGCAAAAGAGAG 2793  
2335 ----- 2334  
2794 TTTTAAGAAAGATACCATTTTGTCTGGAACGCTTGGAAGCAATCATGCAATAG 2853

2335 ----- 2334  
2854 AGCAATTAATGAGGACAAATAAGCCCAATATAGAGTACTGGGCAAAAGAGTAG 2913  
2335 ----- 2334  
2914 GACTGAAGGCTGTGCTTCATAAAACCACAGCTTGAAAGCGCATCAACGGGAAAT 2973  
2371 tggtaactcttccagtcagatcaagaagaaattgactatgataccataccagt 2430  
2974 TCGTACTACTCTTCAGTCAATGATCAAGAGAAATGACTATGATGAATACCAT 3033  
2431 aatgaagaagaaatttgcacattatgaatgaatgaatgaatgaatgaatgaat 2490  
3034 AATGAAGAAAGAAATTTTACATTTATGATGAGATGAATGAATGAATGAATGA 3093  
2491 tcaaaagaacacagacactatttctcgtcagtgagtgagcctcgtgattatgag 2550  
3094 TCAAAAGAAACACGACACTATTTATGCTGCAGTGAAGAGGCTCTGGGATATG 3153  
2551 gagtagctcccaatgcttcaagaacagagctcagatgagtgagtgagtgag 2610  
3154 GAGTACTCTCCCATGATGTTCTAAGAAACAGGGCTCAGAGTGGCATGCTCAG 3213  
2611 gaaagtgtttccagaatctatgagctccttactcaaccccttataccgtgaga 2670  
3214 GAAAGTGTGTTTCCAGGAATTTACTGATGAGCTCTTACTCAGCCCTTATAC 3273  
2671 actaaatgaacatttggagctcctggtggtcataataagaagaaattgaataat 2730  
3274 ACTAATGAACATTTGGAGCTCCTGGGCTATATTAAGAGCAAGATTAATATAT 3333  
2731 catlgaacttcaagaatcagagcctcgtccctatcccttactatcagctat 2790  
3334 CATGCTAATTTTCAAGAAACAGGCTCTGCTCCCTATTCCTTATATGAGCT 3393  
2791 ttatgagaagatcagaagcaagagcagaacctgaaaaaacttgcagaagcaat 2850  
3394 TATGAGGAAGATCAAGGCAAGGACAGACACTAATAAAATCTTGCTCAAGCT 3453  
2851 aacaaaacttacttcttgaagaatgcaacataatgcaacccaactaaagatgatt 2910  
3454 AACCAAACTACTTATTTGGAAGTGCACATCATATGAGCACCACATTAAGATG 3513  
2911 ctgcaagcctggtcttattctctgtatgttgaacctggaagaatgctcagag 2970  
3514 CTGCAAAAGCTGGCTTATTTCTGATGTGACCTGGAAGAGTGTCCACAGGCT 3573  
2971 gattgagccctctgtctgcaacactaaacactgaacccctgctcatgagagaga 3030  
3574 GATTGACCCCTTCTGTGCTGCCACACTAAGACACTGAACCTGCTCATGGAGCA 3633  
3031 gacagtagaagaattgtctgttcttcaacacttcttgaatgaaccaaagctgact 3090  
3634 GACAGTACAGGAATTTGCTGTGTTTCAACCATCTTGAATGAACCAAAAGCTG 3693  
3091 cactgaataatgaagaacactgcaaggtccctgcataatccagaatgaatcc 3150  
3694 CACTGAAATATGGAAGAAAGACTGAGGGCTCCCTCAATATCCAGATGGAATCC 3753  
3151 tttaaagaatctacgttccatgaatgaatgaatgaatgaatgaatgaatgaat 3210  
3754 TTTTAAGGAATTAATGCTCTCATCAATCAATGAGCTCATATATGATACATCT 3813  
3211 cttagtaatgagctcagagacaaagatctcagatgtatcgtcctcagcatgag 3270  
3814 CTTAGTAATGGGCTCAGAGATCAAGAGATTCGATGATGATGAGTGGGAGCAT 3873  
3271 aaacatcatctcatcatctcagtgagacatgctcactgtaagaaaaaagagag 3330  
3874 AAACATCATTTCTATTCATTTGATGAGATGATGCTGATGAGTGAAGAGAGTA 3393  
3331 taaatgagcagcttaacatctcatccaggtgttcttgaagaagtgaatgttaccat 3390











QY 2335 ----- 2334  
Db 2554 TTCTGATGATCCATCACCTGGAGCAATAGACATGATAAAGCCGTCTGGAATGACACA 2613  
QY 2335 ----- 2334  
Db 2614 CTTTCAGGCCACAGCTCCATCAGAGTGGGACATGGTATTTAACCCCTGAGTCAGCGCTCCA 2673  
QY 2335 ----- 2334  
Db 2674 ATTTAGATTAAATGAAAGACTGGGAGACACTGCAGATCCCTCTTGCTTGGAATTAACCACTA 2733  
QY 2335 ----- 2334  
Db 2734 TGGTACTCAGATACCAAAAGAGTGGAAATCCCAAGAGAATCCACGAAAAAAGCAGC 2793  
QY 2335 ----- 2334  
Db 2794 TTTTAAAGAAAAAGATACCATTTTGTCCCTGAAAGCCTTGTAAGCAATGCATGCAATAGC 2853  
QY 2335 ----- 2334  
Db 2854 AGCAATTAATGAGGAGACAAATTAAGCCGAATAGAAATCAGCTGGGCAAGCAAGTAG 2913  
QY 2335 ----- 2370  
Db 2914 GACTGAAGGCTGTGCTCTCAAAAACCCAGCTCTTGAAGCGCCATCAAGGGAATTAAC 2973  
QY 2371 tctgtacctcttcagtcagatcaagaaggaaattgacctatgatatacatatacgttga 2430  
Db 2974 TCGTACTACTCTTCAGTCAGATCAAGAGAAATTTGACTATATGATACCATATACGTTGA 3033  
QY 2431 aatcgaagaaggaaattctgacattatgatlgaagatgaataacacagagcccgagctt 2490  
Db 3034 AATGAAGAAGGAAGATTTTGGACATTTATGATGAGGATGAATAATCAGACCCCGCAGCTT 3093  
QY 2491 tcaaaagaanaacacagacatacttattatgtcagtcagtcgagagctctgagatlatggat 2550  
Db 3094 TCAAAAAGAAAAACAGACACTATTTTATTTGTCAGTGGAGAGCGCTCTGGATTAATGAGGAT 3153  
QY 2551 gattagctcccccacatgtctcaagaanaacagggtcagagtcgagtcgctccctaatga 2610  
Db 3154 GAGTAGTCCCCACATGTCTTAAGAAAGGGCTCAGAGTGGAGTGGCTCCACATTA 3213  
QY 2611 gaaagtctctccaggaatttactgatlgtgtccttactcaagccctataacgctgagga 2670  
Db 3214 GAAAGTGTCTTCCAGGAATTTACTGATGCTCCCTTACACGCCCTTAATACCGTGAGA 3273  
QY 2671 actcaaatgacaatttggagctcctggggccaataaagaagaaagtcgaagataat 2730  
Db 3274 ACTAAATGAACATTTGGGACTCCCTGGGGCCATTAATAAGACAGAAATTTGAAGATAATAT 3333  
QY 2731 catgttaactttaagaataagagctctgtccctattcccttactatctaactatttc 2790  
Db 3334 CATGGTACTTTCAGAATACAGGCTCTCGTCCCTATTCTCTATTTCTATACCTTATTTTC 3393  
QY 2791 ttaatgagaagatcaagaagcaagaagcagaacclagaanaaacttgtcgaagcctaatga 2850  
Db 3394 TTTATGAGGAAGATCAGAGGACAGAGCAGAACCTAGAAAAAATTTGTCAAGCCATAATGA 3453  
QY 2851 aaccacaacttacttcttggaaagtgcacaatcataatgagcccaactaaagaatgattga 2910  
Db 3454 AACCAAAACTTACTTTTGGAAAGTGCAAACATCATATGCGACCCACTAAAGATGAGTTTGA 3513  
QY 2911 ctctcaaaagcttggtctattctctgtatgttgacctggaagaagatgtgcactaagcct 2970  
Db 3514 CTCTCAAAAGCCTGGCTTATTTCTCTGATGTGACCTGAAAAAGTGTGCATCGAGGCT 3573  
QY 2971 gatttgagccctctcgtctcagacactaaacactgaaacctgtcacttgagagacaagt 3030  
Db 3574 GATTGGAGCCCTTCTGTGTCGCCACTAACAACACTGAACCTGCTCAATGGAGAGCAAGT 3633  
QY 3031 gacagtaacaggaattgtcctgttcttcaaccacttctgatagagaccaaaagctgctact 3090

Db 3634 GACACTACAGGAATTTGCTGTGTTTTTCACCATCTTTGATGAGACCAAAAGCTGTACTT 3693  
QY 3091 cactgaaataatggaanaaactcgaaggtcccttgcaataatccaatggaatccac 3150  
Db 3694 CACTGAAATATGGAAGAAACTGAGGGCTCCCTGCAATATCCAGATGGAGATCCAC 3753  
QY 3151 tttaagaagatlatcgtctccatcagatcaatcagatgctacataatgatacactcgtg 3210  
Db 3754 TTTTAAAGAGATATATCGCTTCATGCAATCAATGGCTACATATGATGATACACTCGG 3813  
QY 3211 cttagtaatgtcgaagatcaagaatcgaatglatctgtcctcaagcttggcgaatga 3270  
Db 3814 CTTAGTATGCTCAGATCAAGGATTCGATGTATCTGCTCAGCATGGCAGCAATGA 3873  
QY 3271 aaacatcatcttacttcaatctcagtcgagacatgttcaacttgacgaanaaagaagagta 3330  
Db 3874 AAACATCATTTCTATTTCAATTTTCAGTGGACATGTGTCTCCTAGCAAAAAAGAGAGTA 3393  
QY 3331 laaaatgacactgtacaactcctcatalccagtgcttcttgagacagtcganaatgataccatc 3390  
Db 3934 TAAATGSCACTGTATCAATCTCTATCCAGGCTTTTGAAGACATGGAATGTTACATC 3993  
QY 3391 caaagcttggaatttggcggttgaaatgccttatttggcgagcatctacatgcttggatgag 3450  
Db 3994 CAAGCTGGAATTTGGCGGTGGATGCTTATTTGGCGAGATCTACATGCTGGATGAG 4053  
QY 3451 cacacttctcgtgtgtaagaacaataagtcgagacccctctggaaatgctctctgaca 3510  
Db 4054 CACACTTTTCTGTGTATGATACAAATAGTGTACACTCCCTGGGAATGGCTCTTGACA 4113  
QY 3511 catlagagatttcaagatlaacagcttcagagacaatalgagacagtcggcccaagctgtgc 3570  
Db 4114 CATTAAGATATTTTCAGATTACAGTTCAGACATATGAGACAGTGGGCCCAAAAGCTGGC 4173  
QY 3571 cagacttcaatctccggatcaatcaatgtccttggaagcaaccaagagcccttcttgat 3630  
Db 4174 CAGACTTCATTAATTCGGGATCAATCAATGCTTGAGAGACCAAGAGGCCCTTTCTTGAT 4233  
QY 3631 caaagtgaatcgttggacccaatgatatcatcagagcaatcaagaccaggttccctgtca 3690  
Db 4234 CAAGTGTATCTGTTGGACCAATGATTTATTCAGGCATCAAGACCCAGGTCCTGTCA 4293  
QY 3691 gaagttctcagcctcatalctctcagtttatalcatatglatatgctctgataggaaga 3750  
Db 4294 GAAGTCTCCAGCCTCTCATCTCTCAGTTATCATGATGATGATGATGATGGAAGAA 4353  
QY 3751 gtgagagactatcggaggaatctccactggaaccttaatgtcttcttggcaatgaga 3810  
Db 4354 GTGGCAGACTTATGAGGAATTCACCTGGAACCTTAATGTCTTCTTTGGCAATGTGA 4413  
QY 3811 ttcatctgggataaacaacatactttaaaccctccaattatgtctcgatatacctgtt 3870  
Db 4414 TTTCATCTGGGATTAACCAATATTTTAAACCTTCACATTAATTTGCTGTGATACATCCGTTT 4473  
QY 3871 gcaccaactcatatagcatctcgaagcactcttcgcatgagttgaatggcctgtgattt 3930  
Db 4474 GCACCAACCTCATATATGACATTCGACAGACTCTTCGATGAGTGAATGGGCTGTGATTT 4533  
QY 3931 aaatagttgcagcatcgcattgggaatggagagttaagaacatatcagatgcacagattac 3990  
Db 4534 AAATAGTTGCAGCATGCGCATTTGGGAATGGAAGTAAGCAATATCAATCAACAGATTAC 4593  
QY 3991 tgccttaacttacttcaacatagtttggccaagcttgctcccttcaaaagctcgaactca 4050  
Db 4594 TGCTTCACTCTACTTTTACCAATATATGTTTTCACCTGGTGTCTCTTCAAAAGCTGCACCTTCA 4653  
QY 4051 cctccaaggagagtaatgcttggaagcctcagtcagtcgaataatccaagaagtgtgtca 4110  
Db 4654 CCTCAAGGGAAGATATGCTGTGAGACCTCAGAGTAATTAATCCAAAGAGTGGGTGCA 4713  
QY 4111 agtggactccagaagacaatgaaagtcaagagtaacttaactcaaggagtaaatctct 4170



Db 4714 AGTGACCTCCAGAGACAAATGAAGTCACAGAGTAACACTACAGGAGTAATCTCT 4773  
QY 4171 gcttacacagatgatgtgaaggagttccatctccatccagcaatgaatgccaatgtg 4230  
Db 4774 GCTTACACAGCATGTATGTGAAGGAGTTCATCTCCAGCAGTCMAATGAGCCATCAGTG 4833  
QY 4231 gactctcttttcaagaatgcaagaatgaatgttttcaaggaaatcaagctctcac 4290  
Db 4834 GACTCTCTTTTTCAGATGAGCAAGTAAGGTTTTCAGGGAATGCAAGACTCTTTCAC 4893  
QY 4291 acctgtgtgaactcttagaccaccgttactgaactgactaccttgaatcaaccacca 4350  
Db 4894 ACCTGTGTGTAACCTCTAGACCCACCGTTACTGACTCGCTACCTGGAATTCACCCCA 4953  
QY 4351 gaagtgtgtgacacagatgtccctgaagatgtaggttctgtgctgcaaggcaagacct 4410  
Db 4954 GAGTGTGGTGACACCAAGATGTCCTCGAGATGAGAGTTCGTGGCTGCGAGACAGGACCT 5013  
QY 4411 ctactga 4417  
Db 5014 CTACTGA 5020

RESULT 10  
US-09-243-539-1  
; Sequence 1, Application US/09243539  
; Patent No. 6130203  
; GENERAL INFORMATION:  
; APPLICANT: VOORBERG, Johannes J.  
; TITLE OF INVENTION: HYBRID PROTEINS WITH MODIFIED ACTIVITY  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/243,539  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/558,107  
; FILING DATE: 13-NOV-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: ISACSON, John P.  
; REGISTRATION NUMBER: 33,715  
; REFERENCE/DOCKET NUMBER: 30472/212  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)672-5300  
; TELEFAX: (202)672-5399  
; TELEX: 904136  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 5035 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 35..5017  
; US-09-243-539-1

Query Match 84.8%; Score 3743.8; DB 3; Length 5035;  
Best Local Similarity 87.5%; Pred. No. 0;  
Matches 4365; Conservative 0; Mismatches 52; Indels 570; Gaps 2;

QY 1 catgcaaatagagctctcaactgctcttctctgtgaccttlttgagattcgtcttagtc 60  
Db 34 CATGGAATATAGAGCTCTCCACCTGCTCTCTGTGCTTTTGAGATTGCTTTATGTGC 93  
QY 61 caccagaagaatactaccctgggtgagtggaactgtcaatgggaactataatgaatgtact 120  
Db 94 CACCAGAAGATATACCTGGGTGAGTGAAGTTCATATGGGACTATATGCAAGTGTATCT 153  
QY 121 cggatgagctgtcctgtgtgagcgaagaattcctcctagagtgccaaatccttccatcaa 180  
Db 154 CGGTGACCTGCTGTGTGAGCCCAAGATTTCCTCTAGAGTGCCAAAATCTTTCCATTCAA 213  
QY 181 caactagctgtgtacaaaagaactcgtttgtagaattcaagttcaacttlttcaaat 240  
Db 214 CACTCAGTGTGTACAAAAGACTGTTGTATGAATTCACGATACCTTTTCAACAT 273  
QY 241 cgctaaagcaagcgcaaccctgagatgggtctgtagtgcccaaccacccagctgaggtta 300  
Db 274 CGCTAACCCAAAGGCCACCCCTGGATGGGTCTGTAGTCTTACCATCCAGGCTGAGGTTTA 333  
QY 301 tgatcacgtgtcaatlaacttaagaaacatggtcttccatcctgtcagttctatgtct 360  
Db 334 TGATACAGTGTGATATACCTTAAGAACATGGCTTCCATTCCTGTCAAGTCTATGCTGT 393  
QY 361 tgggtgtactcctactggaagaactctctgagggagctgaaatagatgaatgaaccagtaag 420  
Db 394 TGGTGTATCCTACTGGAACCTTCTGAGGAGCTGGAATATGATGATGACCACTCTCAAG 453  
QY 421 ggaagaagaatagataaagcttccctgtgtggaagcatatgattctcggcaggtcct 480  
Db 454 GGAGAAAGAAATGATATAAGTCTTCCCTGGTGGAGCCATATATTTCTGGCAGGTCT 513  
QY 481 gaaagaatgtgtccaaatgagcctctgaccactgtgtccttaactactacatacttttca 540  
Db 514 GAAAGAAATGTGTCAATGAGCTCTGACCACAGTGTGCTTACCTACATATCTTCTCA 573  
QY 541 tgtgagactgtgtaaaagacttgaattcagagccctcatatggagccctactaatatgtaaga 600  
Db 574 TGTGAGCTGTGTAAGACTTGAATTGAGCTTCATTTGAGGCCCTTACTATATGTAAAGA 633  
QY 601 aggaagctgtgccaagaaagacacagacctgtgacaaattataacttlttgtct 660  
Db 634 AGGAGCTGTGGCCAGGAAAGAACACAGACTTGCACAAATTTATATCTTTTGTGT 693  
QY 661 atttgatgaaggaaagtgtgcaactcagaacaaagaactccttgaatgagtaagga 720  
Db 694 ATTTGATGAAGGAAAGTGTGCACCTCAGAAACAAAGACTCTTGTATGAGATAGGA 753  
QY 721 tgcctgcactgtcctcggcctgtgcttaaatgacacagctaaatgtatgtataacaggtc 780  
Db 754 TGCCTGATGTGCTCGGGCTGTGAAATGACACAGTCAATGTATGTAAACAGGTCT 813  
QY 781 tctgccaagctgtatgtgagccaagaatcaagctatcttggcatgtgtaagtgag 840  
Db 814 TCTGCCAGGTCTGATTTGATGACCAAGGAATCAGTATTTGGCATGTGATTTGAAATGG 873  
QY 841 caccactcctgaagtgtcactcaatlatctcctgaaggtcacacatttctgttgaagacca 900  
Db 874 CACCACCTCTGAAGTGCATCAATATTCCTGGAAGGTGACACATTTCTGTGAGGAACCA 933  
QY 901 tgcgcaaggtctccttggaaatctcgccaataaacttcttactgtgtcaaacactctgtat 960  
Db 934 TCGCCAGGCTGCTCTGTGAATTCGCCAAATCTTTCCTTACTGCTCAACACCTTGTAT 993  
QY 961 ggaacctgtgacagtttctactgttttgaatctctctccaccacaatgagatgagga 1020  
Db 994 GGAACCTTGGACAGTTTCTACTGTTTGTATATCTCTTCCACCAACATATGAGATGGA 1053  
QY 1021 agcttatgtcaaatgtagacagctgttcagaggaagaccacaactagaatgaataatga 1080  
Db 1054 AGCTTATGTCAAAAGTGAAGAGCTGTGTCAGAGGAACCCCACTGAGATGAAAAATATATGA 1113



QY	1081	agaagcggnaagactatgatagaatgacttactgatacttcgaaatggaaatggtgcagtttga	1140
Db	1114	AGAAGCGGAAGACTATGATCATGATATCTTACTGATTCTGGAATGGATGTGGCAGGTTTGA	1173
QY	1141	tgaatgaacttcccttccttccttccaaattgcgtcaagtttgcagaagacaacctaaac	1200
Db	1174	TGATGACMACTCTTCCTTCTCTTTATCCAAATTGCGTCAAGTTGCCAAGAACGCTCTTAAC	1233
QY	1201	tttggtacatlaaatlgtcgtcgaagaggaagcttggactatgtcctctagtccctgc	1260
Db	1234	TTGGGTACATTACATTGCTGCTCTGAAGAGAGACTGGGACTATGCTCCCTTAGCTCCGCG	1293
QY	1261	ccccgatgcagaagaatataaaagtcataatttgaacaatggccctcaagcgatctgtag	1320
Db	1294	CCCCGATGACAGAAAGTTATMAAGTCAATTTTGAAACAATGGCCCTCAGCGAATGGTAG	1353
QY	1321	gaagatacaaaaagctccgatctttagatacacagaatgaaaccttaagctctgtagc	1380
Db	1354	GAAGTACAAAAAGTCCGATTTATGGCATACACGATGAAACCTTTAGACTCGTGAAGC	1413
QY	1381	tattcagcatlgaatcaggaatccttggaccttactttaggggaagtctggaacacact	1440
Db	1414	TATTCCAGCATGATCCAGGAATCTTGGACCTTTACTTTATGGGGAAGTTGGAGACACACT	1473
QY	1441	gtttgatataatlaagaatcaagaacagacaacatataaactcaacctcaaggaatcac	1500
Db	1474	GTTTGATTATATTTTAAAGATACAAACCAACCAACCATATTAACATCTCACCGGAATCAC	1533
QY	1501	tgaatgcgtcccttcttattcaaggagatataccaaaagctgtaaaacatttgaagatt	1560
Db	1534	TGATGTCCTGCTTCTTGTATTCAAGGACATTTACCAAAAGTGTAACATTTGAAGGATTT	1593
QY	1561	tccaattctgcagaagaaatattcaaaatalaaitygcagtgactgtagaagaatggcc	1620
Db	1594	TCCAAATTCTGCAGAGAAATATTCAATPMAATGACAGTGCATGACAGATGGGCC	1653
QY	1621	aactaaatcaagatccctcggtgctgcagccgctatactctagtcttgtaataatggagag	1680
Db	1654	AACATAATCAGATCCCTCGGCTGCAGCCCTATTACTTAGTTCCCTTAATATGGAGAG	1713
QY	1681	agatcagactcaagagactatgtggccctccctcatctgtctacaaagaatctgtagata	1740
Db	1714	AGATCTAGCTTCAGGACTCTATTGGCCCTCTCTCATCTGCTACAAABAATCTGTAGATCA	1773
QY	1741	aagaggaaaaccagataatgtcagacaaagaggaaatgtcatcctgtcttcgtlatltga	1800
Db	1774	AAGAGAAACCAAGATATATGTCAGACAAAGAGAAATGTCATCCTGTTTCTGTATTTGATGA	1833
QY	1801	gaaccgaagctgtgtaccccaagaagaatatacaagcttctccccaatccagcttgaat	1860
Db	1834	GAACCCAACCTGGTACTCTCAACAGATATTAACAAGCTTCTCCCAATCCAGCTGAGAGT	1893
QY	1861	gaagctctggaatccagagttccaaagctccaaacatcatgacaagacatcaatgtctatg	1920
Db	1894	GAAGCTTGAGATCCAGAGTTCCAAAGCTCTCAACATATGACACAGCATCAATGGCTATGT	1953
QY	1921	tttctgaatcttgcagttgtcagtttgttctgtatgcaatggatggcabaactgttacaattcaag	1980
Db	1954	TTTTTGATATTGGCAATTGTGCATTTGTTTGCATAGAGGTGGCAATCGTATCAATTCTAAG	2013
QY	1981	catctgagacacagactgacttcccttctgtcctctctctcggatatactacccttcaaaccaa	2040
Db	2014	CATTTGAGAGACACACATGACTTCTTTCTGTCTTCTCTCTCTGAGATATCCTTTCAAAACCAA	2073
QY	2041	aatgtgcataaagacaacataccctcattcccatctcaaggagaactcttcatgtct	2100
Db	2074	AATGTGCTATGAAAGACACTCACCTATTTCCATTTCTCAGGAGAAACTGTCTTCATGTCT	2133
QY	2101	gattgaaaaaccagctcatatgatatctctgggttgccaacaactcagacttctggacaagag	2160
Db	2134	GATGCAAAACCCAGGCTATGAGATTTCTGGGGTGGCAACACTCAGACTTTGGGAACAGAGG	2193
QY	2161	catgacgccttactgaaggttctctagttgtgacaagaacactgtgtatattacgaga	2220

[illegible]



Db 3274 ACTAATGACATTTGGGACTCTGGGGCCATATATAGACAGAGTGAAGATTAATAT 3333  
OY 2721 catgttaacttcaagaatcagcctctcctcctattccttctatcttaagcctatttc 2790  
Db 3334 CATGTATCTTTCAAGAAATCAGGCTCTGTCCTATTCTCTTAATCTTGCCCTTATTTC 3393  
OY 2791 ttatgagaagatcagagcagagcagaacccatagaanaaactttgcaagcctatga 2850  
Db 3394 TTATAGAGAAATCAGAGGAGAGAGAGAACCTAGAAAACCTTTGTCAAGCCTAATGA 3453  
OY 2851 aaccaaacttacttttggaaagtgcacaatcatalatgagcccaactaaagatga 2910  
Db 3454 AACCAAACTACTCTTTGGAAAGTGCAACATATGGACCCACTTAAGATGATTGGA 3513  
OY 2911 ctgaaagccgggctatttctctgattgttgaccttggaaaaagatgtgcaactg 2970  
Db 3514 CTGCAAGGCTGGGCTTATTTCTGTGATGTGACCTGAAAAAATGTGCACTGAGGCT 3573  
OY 2971 gattggaacccctctgctcctcactaacaactgaacccctgcatgtggaagaagt 3030  
Db 3574 GATTGGAACCCCTCTGCTGCTCCACTAACACACTGACACCTCTCATGGAGACAAAGT 3633  
OY 3031 gacagtaacgaatctgctcctgttttcaacacatcttgatgagaccaaagctgta 3090  
Db 3634 GACAGTACAGGAATTTGCTGTTTACACATCTTGATGAGACCAAAAGCTGTAAT 3693  
OY 3091 cactgaaataatggaagaactgcaagcctcctgcaataatccagttgaaagctccac 3150  
Db 3694 CACTGAAATATGGAAGAACTGCAAGGCTCCCTGCAATATCCAGATGGAAGTCCAC 3753  
OY 3151 tttaaaagagaatatacgtctcctcaatgaatgctacataatgatacaacactg 3210  
Db 3754 TTTTAAAGAGATATGCTTCCATGCAATCAATGAGCTACATATGATACACTG 3813  
OY 3211 cttaagtaatggtcagagtaaaagatctgattctcctcgaactgggcaagatga 3270  
Db 3814 CTGTGATATGCTCAGATCAAAAGATCGATGCTGCTGACATGAGCAATGA 3873  
OY 3271 aaacatcatctatctatctcagtgagatggttcaactgtgaagaaaaagggagta 3330  
Db 3874 AAACATCATCTATTAATTAATTCAGTGACATGCTGTCATGCAAAAAGAGAGTA 3933  
OY 3331 taaaatgacactgacatcctcctacatcaggttcttggagacagtggaatgttaccat 3390  
Db 3934 TAAATGGAAGTATGATCATCTATCCAGAGTGTGAGAGCAATGGAATGTTACATC 3993  
OY 3391 caaagctggaatttggcgggtggaagtgcttatttggcagacatctacatgctggatga 3450  
Db 3994 CAAAGCTGGAATTTGGCGGGTGGAATGCTTATGCGAGCATCTACATGCTGGATGAG 4053  
OY 3451 cagacttttctgtgtacagcaatgaagtgtcagactcccttgggaattgtcttgaca 3510  
Db 4054 CACACTTTTCTGTGTAAGCAATTAAGTGTGACATGCTCCCTGGGAATGGCTTGTGACA 4113  
OY 3511 catagaagatttcaagattcagactcagagcaaatatgagcaagtgggcccaaaagctggc 3570  
Db 4114 CATATAGAGATTTTCAAGATTACAGCTTACAGGACATATGAGACAGTGGGCCCAAGCTG 4173  
OY 3571 cagactcatattcggatcaatcaatgcttggagcacaagaagccttcttgat 3630  
Db 4174 CAGACTTCATTAATCCGATCAATCAATGCTTGAGCACCAAGAGGCCCTTTCTTGAT 4233  
OY 3631 caaagttagactgttggcacaatgattatctcagacataagaagccagggtgcccga 3690  
Db 4234 CAAGGTGATCTGTTGGACCAATGATTATTCAGGCATCAAGGCCAGGGTGCCTGCA 4293  
OY 3691 gaagttcagcctcactacatctcagttatcatcatgtatagttcttgatggaagaa 3750  
Db 4294 GAAATTTCCAGCCTTACATCTCTCAGTTTATCATGATGATGCTTGTATGGAAGAA 4353  
OY 3751 gtggcagacttatcagagaaattccactggaaccttaatggcttcttggcaatgtga 3810  
Db 4354 GTGGCAGACTTATCGAGAAATTCACATGGAACCTTAATGCTTCTTTGGCAATGTGA 4413

OY 3811 ttcaatcgggaataaacaatatltttaaaccctcaattatgtcgtatataatcgttt 3870  
Db 4414 TTCAATCTGGGATTAATAACATATATTTTAACTTCATATTTGCTGATACATCCGTTT 4473  
OY 3871 gcaaccactcatatatacatctcgaacactcttcgaatgtgattgtgtgttt 3930  
Db 4474 GCACCAACTCATATATATGATATGCGACACCTCTCGCATGAGTGAATGGGCTGTGATTT 4533  
OY 3931 aaatagttcagcatgcatgtggaatgagatgaagaatcatalcagatgcagattac 3990  
Db 4534 AAATAGTTGACAGCATGCGCATGGGAATGAGAGTAAGACATATCAATGCACGATTAAC 4593  
OY 3991 tgcctacactcatccaatatgatttgacactgtgtccttcaaaagctgtactca 4050  
Db 4594 TGCTTCACTCACTTACCATATATGTTTGCCACCTGGTCTCTTCAAAAGCTCGACTTCA 4653  
OY 4051 cctcgaaggagagatgactgctgagacactcaagtgaataatcacaagaagtgtgca 4110  
Db 4654 CTTCAAGGGAGAGATATGCTGAGACCTCGAGTGAATATCCAAAGAGTGGCTGCA 4713  
OY 4111 agtggacttccagaagaacaatgaagtcacagagtaactactcagggagtaaatctct 4170  
Db 4714 AGTGACTTCCAGAAAGCAATGAAGTCAAGAGTAACTACTCAGGAGTAAATCTCT 4773  
OY 4171 gcttaccagatgatgtggaagggttcctcatctcgaagatgaagatgccaatg 4230  
Db 4774 GCTTACCAAGCATGTATGTGAAGAGTTCCTCATCTCCAGCATGAATGAGCTGATG 4833  
OY 4231 gactccttlttcaagaatgcaagtaaggttlttcaagggaaatcaagactcctcac 4290  
Db 4834 GACTCTTTTTCAGATATGCAAGTAAGGTTTTCAGGGAATCAAGCTCCTTAC 4893  
OY 4291 aactgtgtgaactcctcagaccacccgttaactgactcgttacttgaattcaacccca 4350  
Db 4894 ACCGTGTGAACTCTAGACCCACCGTTACTACTCGTACCTTGGAATTCACCCCA 4953  
OY 4351 gaattgggtcagacagattccctcagatgagaggttctggctggagcagagagact 4410  
Db 4954 GAGTGGGTGACCAAGATTGGCTTGAGAGATGAGAGTTCGGCTGCGAGGACAGGACT 5013  
OY 4411 ctactga 4417  
Db 5014 CTAATGA 5020

RESULT 11  
US-08-670-707A-38  
; Sequence 38, Application US/08670707A  
; Patent No. 5859204  
; GENERAL INFORMATION:  
; APPLICANT: Lollar, John S.  
; TITLE OF INVENTION: Hybrid Human/Animal Factor VIII  
; NUMBER OF SEQUENCES: 40  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.  
; STREET: 5370 Manhattan Circle Suite 201  
; CITY: Boulder  
; STATE: Colorado  
; COUNTRY: USA  
; ZIP: 80303  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/670,707A  
; FILING DATE: 26-JUN-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: WO PCT/US94/13200  
; FILING DATE: 15-NOV-1994



```

: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/212,133
: FILING DATE: 11-MAR-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/864,004
: FILING DATE: 07-APR-1992
: ATTORNEY/AGENT INFORMATION:
: NAME: Greenlee, Lorance L.
: REGISTRATION NUMBER: 27,894
: REFERENCE/DOCKET NUMBER: 75-95F
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 303/499-8080
: TELEFAX: 303/499-8089
: INFORMATION FOR SEQ ID NO: 38:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 4334 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: not relevant
: MOLECULE TYPE: cDNA to mRNA
: HYPOTHEICAL: NO
: ORIGINAL SOURCE:
: INDIVIDUAL ISOLATE: Factor VIII lacking B domain
:
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 3..4334
: US-08-670-707A-38
```

```

Query Match 72.28; Score 3189.8; DB 2; Length 4334;
Best Local Similarity 83.5%; Pred. NO. 0;
Matches 3692; Conservative 0; Mismatches 637; Indels 90; Gaps 3;
```

```

QY 2 atgcaaatagagcctccacctgctctcttctgctgcttggcgaatttgcctttagtcc 61
DB 3 ATGCAAGCTAGAGCTCTCCACCTGCTGCTTCTTGCTGCTCTTGCCACTCGGCTTTAGTGCC 62
QY 62 accagaagataactactcctggtgctgagtggaactgcatggaactatataatgaaatga---t 118
DB 63 ATCAGGAGATGACTACTGCTGGCGGAGTGAGACTGCTCGGACTACCGGAAAGTAACATC 122
QY 119 ctggtgtagctgctctggtggaagcaagatttccctcctagaatgagcaaatctttccattc 178
DB 123 CTCGCTGAGCTGACAGTGAGCACCAGATTCTCTGCTACAGCGCCGAGACTCTTCCGCTG 182
QY 179 aacacccagtcggtgtacaaaagaacctgtgtgtgagaattcaacaggtccacctttcaac 238
DB 183 GGCCCGTACAGTCTGTACAAAAGACTGTGTTGTAAGTTTCACGGATCACTTTTCAGC 242
QY 239 atcgctaagccaagccaaccttgatgggtctgctaggttccctaccatccaggttgaggt 298
DB 243 GTTGCCAGGCCAGGCCACCATGATGGGTGCTGGGTGCTACCATCAGGCTGAGGTT 302
QY 299 tatgatacagtggtcattacaactaagaacatggtctcccatcctgtaagcttcaagct 358
DB 303 TACACACAGCGGTGCTGTTACCTGAAGAACATGGCTTCTCATCCCTAGTTCACCT 362
QY 359 gtgtgtatctactcagaagcttctcgaaggagctgaattgatatgatacagaaccaagtaa 418
DB 363 GTCGGCTCTCTCTCTGGAATCTTCGAAAGCGCTGAATATGAGGATACCCAGCCNA 422
QY 419 agggagaagaagaatgataaagcttcccttggtggaagccatacataatgctggaagtc 478
DB 423 AGGAGAGAAGAGAGAGATGAAGTCTCTCCGCTAAAGCCAAACCTAGCTGCGAGAGTC 482
QY 479 ctgaagaagaatggttccaatgagctctgaaccaatgagcttaccctactacatatattct 538
DB 483 CTGAAAATAAATGCTCCAAACAGCTCTGACCCACCATGTCTACACTACTCATACCTGCT 542
QY 539 catgtgacactggttaaaagacttgaattcaggccctcatctgtagccctactagtagtaga 598
DB 543 CACGTGACCTGgtGAAGAGACTGAATTCGGGCTCATATGGAGCCCTGCTGTTGTAGA 602
```

```

QY 599 gaaggagctcgtccaaagaagaacacagacacttcacaaattatactacttcttgc 658
DB 603 GAAGGAGCTGTACCAAGAGAAAGACCCAGAACCTGACGAAATTTGTACTCTTTTGGC 662
QY 659 gtaattgatgaagggaagaagtgtgacatccagaagaacatcccttgatgagataggt 718
DB 663 GCTTTGATGAAGGAAAAAGTTGGCCTCACCAAAATGACTCTCTGACACGGGGCATG 722
QY 719 gatgtcgaatcgtccggtgctgaaatcgaatgaatgacacagtcgaatgttatgaaacag 778
DB 723 GATCCCGCACCTCGCAGGGCCCGCCAGCTCATGACACAGTCATGCTATGTAACAGG 782
QY 779 tctctgcaagtlctgattgtagtgcacagaagaatcgaatcgaatgtagtgaatg 838
DB 783 TCTGTCCAGAGCTGATGAGATGATGAAGAAATGACTGCTACTGCAAGTATGGAATG 842
QY 839 ggcacactccctgaagtgtcactcaatatctctggaaggtccacaaattctgtgaggaac 898
DB 843 GGCACACCCCGGAAGTGCACCTCAATTTTCTTGAAAGGCCACAGGTTCTCGTAGGCGAC 902
QY 899 catgccaaggcgtccttgaaatctcgcaaatccttctactgctcaaacactctg 958
DB 903 CATGCCAGGCTCTCTTGGAAGTCTCGCCACTAATCTTCCCTACCTGCTACAGCATTTCTG 962
QY 959 atgacacttggaacagtlctactglttltgcatatactcttcccaacacatgtagcatg 1018
DB 963 ATGAGACTTTGGCAGTCTACTGTTTGTGATATCTCTTCCACCAACATGTTGGCATG 1022
QY 1019 gaagcttatgtcaaatgatacagctgtgccagaaggaacccaactagaatgaaataaat 1078
DB 1023 GAGGCTCAGCGTCAGAGTGAAGAGCTGCGCCGAGAGAGCCCGCAGCGAGAAAGCGAT 1082
QY 1079 gaagaagcgggaagactatgatatgatatcttacttcgaatgagatggtgagtgagttt 1138
DB 1083 GAAAG---GGAAGATTATGATGACATATTTGTACANCTGACATGACATGCTGCGGCTC 1139
QY 1139 gatgatacaactcctcctcttataccaaatcgcgtgaattgccaagaagacatctaa 1198
DB 1140 GATGCTGACAGAGTGTCTCCCTTATCCAAATCCGCTGGTGGCAAGAAAGCATCCCAA 1199
QY 1199 acttggtacattatattgctgctcgtgaagaaggaagacatggtctatgcttccctagctc 1258
DB 1200 ACCTGGGTGCACTACATATCTGCAAGAGGAGAGACTGCGCTGCGCCCGCGCTGCC 1259
QY 1259 gcccccgaatgacaagatlaaagaatcaatataatgaaatgacccctcagcgatggt 1318
DB 1260 AGCCCACTGACAGAGATTTAAAGTCTCTACTTGAACATGCTCTCCACGCAATGGT 1319
QY 1319 aggaagtaacaaaagctcgaattatggtacatacagaatgaaccttgaagctcgtgaa 1378
DB 1320 AGGAATATCAAAAAAGCTCGATTGCTGCTCTTACACGATGTAAACATTGAAGACTCGTAAA 1379
QY 1379 gctattcagcatgaatcagaatcttgggaaccttactctttagtggaagtgtgagaaca 1438
DB 1380 GCTATTTCCGTATGATTCGGAATCCTGGGACCTTTACTTTATGGAAGATGGAGACA 1439
QY 1439 ctgttgatataatlaaagaatcaagaagcagacacataaactcaccctcaagatc 1498
DB 1440 CTTTGTATATATTATGAATTAAGCCAGCGACCATATTAACATCTACCTCATGAGATC 1499
QY 1499 actgattgcctccttglttattcaagagatatacacaagaagtgtaaaacatlttgaaagt 1558
DB 1500 ACTGATGTACAGCGCTTTCACCCAGGAGACTTCTAAAGGTTGGAACAATTGGAAGAG 1559
QY 1559 ttccaattctgccaaggaatatatcaaatataatgagacagtgtatgagaatagg 1618
DB 1560 ATGCCAATTTCCAGGAGAGACTTCAAGTATTAAGAGAGTGTGGAAGATGGG 1619
QY 1619 ccaactaaatagatccctcgtgtcctgaaccgcctattacatgattcgttaatatgag 1678
DB 1620 CCAACCAAGTCCGATCTCGGTGCTGACCCGCTACTACTGCACTCCATTAATCTAGAG 1679
QY 1679 agagatctagctcagagactatgtgcccctctctctcatctgtctacaaagaatctgtagat 1738
```



Db 1680 AAAGATCTGGCTTGGGACACTATGTGGCCCTCTCCTCATCTGCTACAAAGAAATCTGTAGAC 1739  
Qy 1739 caaagagaagaaaccagataaigtacagacaagaagaaigtcatcctggtttcttgat 1798  
Db 1740 CAAAGAGAAAACCGAGTGAATGTGCAGACAGAAACGTCTCCGTTTCTGTATTCGAT 1799  
Qy 1799 gagaacgaagctgtgtaccccaagataataacagctttctcccaatccagctgga 1858  
Db 1800 GAGAAATAAAGCTGTGTCCTCGCAGAGAAATTTAGCGCTTCTCCCAATCCGATGGA 1859  
Qy 1859 gtgcagcttgagatgccagagttccaaagctccaacatcgtcacaacatcaatggtcat 1918  
Db 1860 TTACAGCCCCAGATCCAGATTCCAAGCTTCTCAACATCATGACACACATCAAGGCTAT 1919  
Qy 1919 gtttttgatagttgcagttgtcagttgttttgatagaagtggtgacacgtgtacatcta 1978  
Db 1920 GTTTTTCATAGCTTGACGCTGCTCGTTTGTTCACAGAGGTGGCATCTGTATCATCTTA 1979  
Qy 1979 agcatgtgagacagactgtactctctctgtctctctctgtgataacctcaaacac 2038  
Db 1980 AGTGTGGAGCAGACAGAGACTTCTCTCCGTCTTCTGTGCTACACCTTCATAACAC 2039  
Qy 2039 aaaatgtctatagaagacacactcacctatcccatctcagagaagaactgtctcatg 2098  
Db 2040 AAAATGCTATGAAAGACACACTCACCTGTCCCTTCTCAGAGAAACGGTCTCATG 2099  
Qy 2099 tcgattgaaaaaccagagttcctatgtatctgtggtggtccacaactcagaacttcggaaca 2158  
Db 2100 TCAATGGAAAAACCGAGTCTCTGGGTCTAGGGGTCACACACTGACCTTGGCGAACA 2159  
Qy 2159 ggcatagccgccttactgaaggtttctagtgtgacagaagaacctgtgtatatacag 2218  
Db 2160 GGGATGACAGCCTTACTGAAGGTTATAGTTGTGACAGGNCATTTGGTTATTTATGAC 2219  
Qy 2219 gacagtataagaatatttcagacatactgtctgagtaaaaaacaatgcatatgaacaaga 2278  
Db 2220 AACACTATGAAAGATTTTCCAGGCTTCTGTGATGGAAGAAATGTCATTGAAACCGAGA 2279  
Qy 2279 agcttctccagaattcaagaacacctagacataagcaaaaagcaatttaatgccaacca 2338  
Db 2280 ----- 2279  
Qy 2339 ccaagtttgaacgcacatcaacggaataactgcgtactctctcagtcagatcaagag 2398  
Db 2280 -----GACATTAAGCCTTCTCTTCTTTCAGCCGGAGAGAGAC 2315  
Qy 2399 gaaattgactatgtatatacatatcagttgaaatgaagaagaagaatttgacattat 2458  
Db 2316 AAATGAGACTATGATGATATCTTCTCAACTGAAGCAAGGAGAGAAATTTGACATTAC 2375  
Qy 2459 gatgagagatgaataacagagcccccgagcttccaagaagaagaacagacatattatc 2518  
Db 2376 GGTGAGAGATAAATCAAGACCTCGACGTTTCAGAAAGAACCCCACTGATTTTCATT 2435  
Qy 2519 ggtcgaattgagaaggtctgtggaattatgagatgtagctcccccacatgftctcaagaac 2578  
Db 2436 GCTGCGGTGAGACAGCTCTGGGATTAAGSAGTACGCAATCCCCCGGGCGCTTAAGAAC 2495  
Qy 2579 agggctgaagtgtgcagtgctccctcagttcaagaagaattgttttccaggaattacag 2638  
Db 2496 AGGGCTCAGAAAGAGAGGTGCCGCTCGGTCAAGAGGTGTCTTCGGGAATTTGCTGAC 2555  
Qy 2639 ggtctcttctacagccctataacacgtggaagaactaaatgaacatttggactccgtggg 2698  
Db 2556 GGGCTCTTCACGAGCGCGTGTCCGGGGAACTCAAAACCACTTGGGGCTTCTTGGA 2615  
Qy 2699 ccataataagagcaagaagtgtgaagataatatacatgtgtaacttcagaaatcagagctct 2758  
Db 2616 CCCTACATCAGAGCGGAAGTTGAAGACATCATGTGTTTCAAAAAACAGGCGCTCT 2675  
Qy 2759 cgtccattctctctatcttaagcctattctatagagaagaatcagaagcagaagaca 2818  
|||||

Db 2676 CGTCCCTATTCTTCTACTAGACCTTATTCTTATCCGATGATCAGAGCAAGGGCA 2735  
Qy 2819 gaacctagaaaaaactttgtcagagcctaatgaaaccaaacttacttttgaaagtcaa 2878  
Db 2736 GAACCTTGACACAACTTGTGTCACAGCCAAATGAACCGAACTTACTTTTGAAAGTCCAG 2795  
Qy 2879 calcatatggcaccactaaagatgttagctgcaagaacgtgggtctttctctgat 2938  
Db 2796 CATCACATGGCACCACAGAAAGACGATTTGACTGCAAAAGCTTGCGCTTCTTGAT 2855  
Qy 2939 gttgaccttgaaaaaagatgtgcactcagagcctgatltgagccctctgtctgccaaact 2998  
Db 2856 GTTGACCTGGAAAAAGATGTGACATCAGGCTTGATGCGCCCTTGATGCTGCGCGCC 2915  
Qy 2999 aacacactgaacctgtctcatgtggaagaagtgacagtagaagaattgtctgttttc 3058  
Db 2916 AACACCTGGAACGCTGTCTCAGGTAGCAAGTGCACCTGCAAGAAATTTGCTGTTC 2975  
Qy 3059 accactttgatgagaccaaagctgtgtacttcaacttgaaatatggaagaagaactgag 3118  
Db 2976 ACTATTTTATGAGACAAAGAGCTGTACTTACTATAATGTGGAAAGAACTGCGCG 3035  
Qy 3119 gctccctgcataatccagatggaagatcccaactttaaagagatlatcgtctccatgca 3178  
Db 3036 GCCCCCTGCACTGACAGATGGAAGACCCACCTGTGAAGAAAAACATTCCTTCATGCA 3095  
Qy 3179 atcaatgtctacataatggaatacactcctggtctagttatgtgtcagaatcgaagatc 3238  
Db 3096 ATCAATGTCTATGATGAGATACACTCCGTGCTTAAATGTGCTCAAGATCAAAAGATC 3155  
Qy 3239 cgaatgtatctgtcagcagctgagcagaatgaaacatccatctatcttcaagtgga 3298  
Db 3156 CGATGTATCTGCTCAGCATGGGACAGCAATGAATAATTCATTTCGATTCATTTTGAAGCGGA 3215  
Qy 3299 catgtgtcactgtlacgaaaaaagagagtalaaaaatgagcactgtgtaaaatctctatcca 3358  
Db 3216 CAGCTGTTCAGTGAAGCAAAAGAGAGAGATTAATAATGSCCGTTCATATCTATCCG 3275  
Qy 3359 ggtgttttggagaagtggaatgttacatccaaagcttgaatttggcgtgtggaatgc 3418  
Db 3276 GGTGTCTTTAGACAGTGGAAATGCTACCGTCCAAATTTGGAATTTGGCAATAAGAAATGC 3335  
Qy 3419 ctatttggcagacatctacatgtcttggatgagcacactttcttctgtgtcaagcaatag 3478  
Db 3336 CTGATTTGCGCAGACACTGCAAGCTGGGATGACACAGACTTCTCTGTGCTACAGCAAGGAG 3395  
Qy 3479 tgtcagaactcccttgggaatgtgtctctgacacattagaagatttcagatlaacgttca 3538  
Db 3396 TGTACAGCTCCACTGGGAATGGCTTGTGACGCAATTGAGATTTTTCAGATCACAGCTTCA 3455  
Qy 3539 ggaacataatggaagtggtgcccccaagctgtgccagacttcaatattccggatcaatcaat 3598  
Db 3456 GGAAGATGAGACAGTGGGCCCAAAAGCTGGCAGACTTCAATTAATTCGGATCAATCAAT 3515  
Qy 3599 ggtctgagacacaaagggagcccttctcttgatcaagtgatctgtgtgaccaaagatc 3658  
Db 3516 GCGTGAAGCAACGAAGATTCGCCACTCTGSGATCAAGTGGATCTGTTGGACCAATGATC 3575  
Qy 3659 attcaagycatcaagaaccaggggtgcccgtcagaagttctcagacctctacatctcag 3718  
Db 3576 ATTACAGGCAATCATGACCCAGGGTGCCCGCAGAAAGTTTTCAGGCTTCATCTCCAG 3635  
Qy 3719 ttatacatcatgtatagttcttgaatggaagaagtgtgcagaactcatcgaagaaatccact 3778  
Db 3636 TTTATCATCATGTACAGTCTTGAAGGAGAACCTGACAGATTCACGAGGGAATTTCCAGC 3695  
Qy 3779 ggaaccttaatgtctctcttcttgcaatgtggaatctcgtggaataaacaatatctt 3838  
Db 3696 GGCACCTTAATGCTCTTCTTGGCAATGTGACGCAATCTGGGATTAACCAATATTTT 3755  
Qy 3839 aacctccaattatgtctcgatatacatcgttttgaccaacatcatatagcatctcagc 3898  
Db 3756 AACCTTCGATGTGGTGTGCTGATCCGTTTGCACCCAAACATTAACAGCATCCGAGC 3815  
|||||



OY	3899	actcttcgcgaatggagtgatgtagggctgtgatcttaataagtctgcagcatcgttggaa	3958
Db	3816	ACTCTTCGCATGAGTGTGAATGGCGCTGTGATTTAACAGTTTGACATGCCCCCTGGAA	3875
OY	3959	gagagtaaaggcaalacacagatgcacagattactgtctcaaccttaaccataatgttt	4018
Db	3876	CAGAATAAAGCGCATCATTACAGACTCACAGATCAACGCCCTCTCCACCCTAACGAATATATT	3935
OY	4019	gccacactgtctccctccaagaagctcgactctcaacctcccagaaggagaglaatgcttgaga	4078
Db	3936	GCCACTGTGTCTCTTTCACAAGCCCAGCTTCACTCCAGG66CGAGAAATGCTTGCGGA	3995
OY	4079	cctcaggTgaataatccaaaaggTtgctgtcgaagtggacttccagaagacaatatgaatc	4138
Db	3996	CCCCGGGTGACGACGCCAGAGAGTGGCTCAGSTGTGACCTGCAGAAAGCGTGAAGTTC	4055
OY	4139	aagaagatgaactaactcgaaggagtaaaatctctgtcttaccagaacatgtatgtgaagagttc	4198
Db	4056	ACAGGCATCACCCACAGGCGCGTGAAMTCCCTGTCAGCAGCATGTATGTGAAGAAATTC	4115
OY	4199	ccaatctccagcagtcgaagaatggcaccatcagTtgactctcttttccgaatggaaagta	4258
Db	4116	CTCGTGTCCAGTGTGAAGAGACGGCCGCGCTGGAACCTGTTTTCTTAGAGACGCCACACG	4175
OY	4259	aaggTtttccaaggaaatcaagaactccctcaacactgtgltgaactctctagaccacaocg	4318
Db	4176	AAGGTTTTTCAGGCGCATCAGGACTCCTCCACCCCGCTGTTGAACGCTTGAGACCCCCCG	4235
OY	4319	ttactpacctcgcactaccttgaaatcacccccagaagtgtggctgcacagatggcctgag	4378
Db	4236	CTGTTTCACCGCGTCACTTGAGGATTCACCCACGACGACTGGCGCAGCATGCCCCCTGAGG	4295
OY	4379	alcgaagttcttggcgtcgcagacagacagactctaactga	4417
Db	4296	CTCGAGGTTCTAGATGTGAGGCACAGGATCTCTACTGA	4334
 RESULT 12 US-09-037-601-38 Sequence 38, Application US/09037601 Patent No. 6180371 GENERAL INFORMATION: APPLICANT: Ioliar, John S. TITLE OF INVENTION: Hybrid Human/Animal Factor VIII NUMBER OF SEQUENCES: 40 CORRESPONDENCE ADDRESS: ADDRESSEE: Greenlee, Winner and Sullivan, P.C. STREET: 5370 Manhattan Circle Suite 201 CITY: Boulder STATE: Colorado COUNTRY: USA ZIP: 80303 COMPUTER READABLE FORM: MEDIUM TYPE: Floppy disk COMPUTER: IBM PC compatible OPERATING SYSTEM: PC-DOS/MS-DOS SOFTWARE: Patent Release #1.0, Version #1.30 CURRENT APPLICATION DATA: APPLICATION NUMBER: US/09/037, 601 FILING DATE: 26-JUN-1996 CLASSIFICATION: PRIOR APPLICATION DATA: APPLICATION NUMBER: NO PCT/US94/13200 FILING DATE: 15-NOV-1994 PRIOR APPLICATION DATA: APPLICATION NUMBER: US 08/212,133 FILING DATE: 11-MAR-1994 PRIOR APPLICATION DATA: APPLICATION NUMBER: US 07/864,004 FILING DATE: 07-APR-1992 ATTORNEY/AGENT INFORMATION: NAME: Ferber, Donna M.			

```

?      REGISTRATION NUMBER: 33, 878
?      REFERENCE/DOCKET NUMBER: 75-95F
?      TELECOMMUNICATION INFORMATION:
?      TELEPHONE: 303/499-8080
?      TELEFAX: 303/499-8089
?      INFORMATION FOR SEQ ID NO: 38:
?      SEQUENCE CHARACTERISTICS:
?      LENGTH: 4334 base pairs
?      TYPE: nucleic acid
?      STRANDEDNESS: double
?      TOPOLOGY: not relevant
?      MOLECULE TYPE: cDNA TO mRNA
?      HYPOTHETICAL: NO
?      ORIGINAL SOURCE:
?      INDIVIDUAL ISOLATE: Factor VIII lacking B domain
?      FEATURE:
?      NAME/KEY: CDS
?      LOCATION: 3..4334
?      US-09-037-601-38

```

Query Match	72.2%;	Score 3189.8;	DB 4,	Length 4334;
Best Local Similarity	83.5%;	Pred. No. 0;		
Matches 3692;	Conservative	0;	Mismatches 637;	Indels 90;
				Gaps 3;

OY	2	accaaatagagctccacccgctctcttcctgagccctttagagctctgcttagagcc	61
Db	3	ATGCACGTAGAGCTCTCCACCCTGTCTCTTTGTGTCTCTTGCACCTCGGCTTTAGTGC	62
OY	62	accagaagatactaccctggtgctgagtggaactgcatacgtgactatagtcaagtga---	118
Db	63	ATCAGAGATACACTACCTGGGCGCATGTGAACCTGCTGGAGTACCGCAAGTGAATCTC	122
OY	119	ctcgtgagctgcctgctgtagcgcgaagatttcctctcctagatgagcaaaacttltccatc	178
Db	123	CTCCGTGAGCTGAGCTGGAGACACAGATTTCTGTCTACACGGCAGAGACTTTCCGTTG	182
OY	179	aaacacctcgtgtgacaaaagactctgtgttgaaatcaacgttcaacttctaac	238
Db	183	GGCCCGTCACTCTGTACAAAAGACTGTGTCTGTAGTTCAGGATCAACTTTTCAGC	244
OY	239	atcgtctaaagccaggccaccctgtagtgcctagtctcctacatccacgcctgaggt	298
Db	243	GTTGCCAGGCCACAGCCACCATGATGATGGGTCTGCTGGTCTTACATCCAGGCTAGATT	302
OY	299	tatgatacagttggttcattacacttaagaanaatggttccacatccgtcaagttcaagct	358
Db	303	TACGACACGCTGCTGTTACCTGTAAGAACAATGGCTTCTATCCGCTTAGTCTTCAGCT	362
OY	359	gttcgtgctactctactggaagaactctcgaaggaactgaaatgatgatcaagcaagtcaa	418
Db	363	GTCGGGCTCTCTCTCGGAATCTTCCGAAGGCGCTGAATATGAGATACACACGCA	422
OY	419	aggagaaagaagatgataaagctctccctcgtgtggaagcatatatalatctgcagctc	478
Db	423	AGGGAAAGGAAGACGATTAAGTCTTCCGCTTAAAGCCAAACCTACCTCTGGCAGGTC	482
OY	479	ctgaagaagatggtcccaatgacctctgaagcccaatgacctctacactcaatcttctc	538
Db	483	CTGAAGAAGAAATGTCTCAACACCTCTGACCCACATGTCTACCTACTCATCTGTCT	542
OY	539	catgtgacctgtglaaagaacttgaattacacagccctcatctgagccctacatgatlgtaga	598
Db	543	CACGTGACCTGTTGAAGACTGTAATTCGGGCTCATTTGGAACCTCTGCTGTTGTAGA	602
OY	599	gaaggagatctggtccaaggaagaacacagacactctgcaaaattatactactcttggct	658
Db	603	GAAGGAGTCTGACCAAGAAAGGACCCAGACCTGTGACAAATTTGTACTCTTTTGGCT	662
OY	659	gatttgtatgaaggaanaaatttgacatcagaanaaagaagactcttgaatcagatagag	718
Db	663	GCTCTTGATGAAGGAAAGTTGGCACTACGCAAGAATGACTCTCTGGACACGGGCCATYG	722



QY	719	gagtcgtcatctcgtcgtggccctggcctaataatgcacacagtcataatggttattgtaacagg	778
Db	723	GATCCCGCACCTTCTCCAGGGGCCAGCCTGGCAATGCGACACAGTCAAGGCTATGATGCACAGG	782
QY	779	tctctgcagctctgattgattgacacaggaataacgtctattgtgcattgattggagatg	838
Db	783	TCTCTGCCAAGTCTGATCGGATGTCTATAAGAAATACAGTCTACTGGACAGTATTTGGATG	842
QY	839	ggcacactctctgaatgtcacctcaataatactccctcgaagtcacacactctctgttaggaac	898
Db	843	GGCACCGACGCCGGAAGTGCATCTCATTTTCTTTGAAGCGCACAGCTTCTCTGTAGGCAC	902
QY	899	cacgcacagcgctcctctggaatctgcocataactcttcctactctgtcacaacactctg	958
Db	903	CATCGCCAGGCTCTCTTGGAGATCTCGCCACTCAACTTTCCTACAGTGCACAGCAATTCCTG	962
QY	959	atggacctggagaagtttactatgittgttaactctctccacacacatgagtcagt	1018
Db	963	ATGGACCTTGGCCAGTCTCTACTCGTTTGTCTATATCTCTCCACACCAATGAGGCGATG	1022
QY	1019	gaagcttatgtcacaagttagacacgctgtccacagaggaacccacactacgatatgaataat	1078
Db	1023	GAGGCTCAGCTCAGATATAAAGCTGCCGCCAGAGGCCACAGCTCGSAGGAAGAAGCTGAT	1082
QY	1079	gaagaagcggaaactatgatagtacttactgattctgaatatgatatgtygttcagttt	1138
Db	1083	GAAAGA--GGAAATTTATGATGACAAATTTGTACGACTCGGACATGAGAGTGTGGTGGCTC	1139
QY	1139	gatgatgaacactctctctctcttcttcacaaatctgcttaagttgtcagaagaacatctctaa	1198
Db	1140	GATGGTGACACGCTGTCTCCCTTTATTCCAAAATCCGCTGGTGTGCCAAACAAATCCCAAA	1199
QY	1199	acttgggtacattacattgctgtcgtcaagaagagagactggaactatgctcccttaagctc	1258
Db	1200	ACCTGGGTGACTACTACTCTCTCGAAGAGAGAGAGACTGGAGCTACGCCGCCGGTCCCTCC	1259
QY	1259	ggccccgcatgacagaagtataaaatgaatcttaattgaaacaatgycctcagcggattggt	1318
Db	1260	AGGCCCAAGTGCAGAGATTTATAAAGTCTCTACTTGAACAGTGGTCHCTCAGCAGAAATGGT	1319
QY	1319	agggaagtaaaaaaagtctcgattattatgcatcacaaagttgaaccttaagacctcgaa	1378
Db	1320	AGGAATATCAAAAACCTCGATTCTGCTGTACACGGATGTATACATTTAAACATCGTAA	1379
QY	1379	gctattcaagcatbaactcaggaactcttggagaccttaactctcttgsgaagttgtagacaa	1438
Db	1380	GCTATTTCCGATGAAATCAGGAATCTCTGGACCTTTACTTTATGGAGAAAGTTGGAGACACA	1439
QY	1439	cgtgtgattatattgaagaatcaagcaagcagacacataataacatctaacctcacaagatc	1498
Db	1440	CTTTTGTATTATTTTAAGATTAAGAGGAGACCCAGCATATTAACATCTACCTCATGGAATC	1499
QY	1499	actgatgtccgtctcttgtattatccaagagattaccaaaagttgtaaaaactttgaagat	1558
Db	1500	ACTGATGTCAAGCCCTTTGACCCAGGAGAACTTTAAAAGGTGGAAACATTTAAAAGAC	1559
QY	1559	tttccaattctgcagagagaataattacaataataatggaacgtatgtagaagatggg	1618
Db	1560	ATGCGCAATTTGCGCAGAGAGAACTTTCAAGTATAATGGACAGTCACTGTGGAAAGATGGG	1619
QY	1619	ccaaactaaatcagatccctcgatgctgcacccgctataactcctagtttcgtataatagag	1678
Db	1620	CCAACCAAGTCCGATCTCTGGTGGCCGAGACCCGCTACTCTGAGGCTCATTAATCTAGAG	1679
QY	1679	agagatcttagcttcagagactcaattggccctctccctcaactcgtctacaagaatctgtatg	1738
Db	1680	AAAGATCTGGCTTTCGGGAATCATTTGGCCCTCTCCATCTGCTCAAAAGAAATCTGTAGAC	1739
QY	1739	caagaaggaacaacagataatgtcaacacagagaaatgtcacctgttttctgtattggat	1798
Db	1740	CAAGAGGAAACCAAGATGATGTACAGACAGAGAAACGTCATCTCTGTTTCTGTATTCGAT	1799
QY	1799	gagaaacgaagctgtgactctcacaagaaatatacacaagcttctctcccaactccagcttga	1858

Db	1800	GAGATCAAAACCTGGTACCTTCGCAGAGAAATTTTCACGGCTTCCTCCCAATCCGATGGA	1859
OY	1859	gtgcagcttggagatgccagattccaaagcctccaaatcatatgcaagcaatgaatgcat	1918
Db	1860	TTACAGCCCCAGGATCCAGAGTTCCAAAGCTTTCATAATATGACACAGCATCAATGGCTAT	1919
OY	1919	gtttctgatagtttgcagatttgtcagatttgtttgcataagatggcatatgcgtatccta	1978
Db	1920	GTTTTTGTATACCTTGACGCTGTGGTTTGTTTGTGCACGAGGTGGCATACGTGATCATTTCTA	1979
OY	1979	agcaattggagcaagactgcacttccttctcgtcttctctctcgtgatatataccttcaaac	2038
Db	1980	AGTGTGGAGCAGACAGAGGACTTCCTCCGTTCTCTTGCGCTACACTTTCAAACAC	2039
OY	2039	aaatgtgtctatggagaacaaacaaacccatattcccatcttcaagggaatcttctcgt	2099
Db	2040	AAAAATGGCTATGTAGAACACACTACCTCTGTCCCTTCACGGAAGAAACGGCTTTCATG	2099
OY	2099	tcgaatggaaaaacccaagttctatgtatctctgggtgtgccaaactcagaactttggagaaga	2158
Db	2100	TCATATGAAACCCAGGTTCTGTGGTCTATGGGTGCGACACACTCAGACTTGGGAAACGA	2159
OY	2159	ggcaatgacgccttactgtaaagtttctagtttgttgacaagaacaactgtgtatattacag	2218
Db	2160	GGGATGACAGGCTTACGTGAAGGGTGTATAGTTGTGACAGGAGACATTGCTATTTATTATAC	2219
OY	2219	gacagttatggaaatatttatagaactctgtgtatgtatgaataaacaatgacattgaaccaaga	2278
Db	2220	AACCTTATGAGATTTTCCAGGGCTTCTGTCTGTAGTGAAGAAATGTCAATTAACCCAGA	2279
OY	2279	agctctcccaagaattcaagaacccctagactagcaaaagaatllaatgtccaccaca	2338
Db	2280	-----	2279
OY	2339	ccagltctgaaacgcacatcaacgaggaaataactcgtactactcttcaagtcagaatcaag	2398
Db	2280	-----GACATPAAAGCTTCTCACTTTCACGCGGAGGAGAAAC	2315
OY	2399	gaattgtaacttgaatgataccataatgaattggaatgaagaagaagaattttacattat	2458
Db	2316	AAAAATGGCTATGTAGTATTTCTTCTCAACTCGAAACAGAGGAGGAAGAAATTTTATCAATTTAC	2375
OY	2459	gataagatatgaataatcagagcccccgcagcttccaaagaagaaacacgacactattat	2518
Db	2376	GGTGAGGATGAATATCAGACACCTCGAGCTTTCACAAAGAAACCGACATATTTCATT	2435
OY	2519	gctgcagatggagaggtctcgtggaattatggtatgtagtagcttcccacatgttctaaagaac	2578
Db	2436	GCTCGGGGAGCAGCTCTGGGATTTACGGGATGACGGAATCCCCCGGGCGCTTAAGAAAC	2495
OY	2579	aggagctaaagatgtagcagtgctcctcaagttcaagaagaattggttttccagaatattcat	2638
Db	2496	AGGGCTTCAGAACGGAGAGTGGCTTCGGTTCTAAAGAGTGGTCTTCGGGAAATTTCTGAC	2555
OY	2639	ggctcccttaccagcccttataacgctgtgagaactaaatgaaacatttggagactccctggg	2698
Db	2556	GGCTCTTTCACGCAAGCGGTGTATCCGGGGGAGACTCAACAAACACTTGGGGCTTTGGGA	2615
OY	2699	ccatataaagagcaagaattgaaagataatcaatggttaacttccagaatcaggctct	2738
Db	2616	CCCTTACATCAAGAGCGGAAGTTGAAGAACATCATGTGTAACCTTTAAAAAACACAGCGCTCT	2675
OY	2739	cgctccattctctctatcttcaagcctatttcttctatgaaagaagatcagaggaagaaga	2818
Db	2676	CGTCCCTATTTCCTTACTGACCTTATTTCTTATTCGGAATGATCAGGACGAAGGGGCA	2735
OY	2819	gaacctagaaaaaacttgttcaagcctaatgaaacaaacacttacttttggaaatgtgca	2878
Db	2736	GAACTTGACACAACTTCGTGTCAGCCAAATGAAGAACAGAACTTACTTTTGGAAATGCGAG	2795
OY	2879	catcataatggacaaccaataaagaatgatttgcactgcaaaagccttggctatttctctgat	2938



Db 2796 CATTACATGGCACCACAGAGAGATTGACAGCAAGCTGGGCTACTTTTCTGAT 2855  
QY 2939 gttagctggaagaaagatgtagctcaagcctgtagttagaccctctgctcgcacat 2998  
Db 2856 GTTACCTGGAAAAAGATGTGCACCTCAGGCTTGATGGGCCCTCTGTATCTGCCGCGC 2915  
QY 2999 aacacacgaacctgctcagatgagagacagtagcagtagcaggaattgctctgttttc 3058  
Db 2916 AACACCTGTAAAGCTGCTCAGCGTAGACAGTACCGTGAAGAAATTTGCTGTTTTTC 2975  
QY 3059 accacttttagtagagcaaaagctgtagcttagcttagaataatgaaagaaagcag 3118  
Db 2976 ACTATTTTGTAGACAAAAGCTGTACTTCTACTGAAATGTGAAAGAACTGCCG 3035  
QY 3119 gtccctgcgaatataccagatggaagatccactttaagaagaatatacgctccatgca 3178  
Db 3036 GCCCCCTGCCACCTGCAGATGGAGAGACCCCACTGTGAAGAAACTATCGCTTCATGCA 3095  
QY 3179 atcaatggtacataatgagatgacactacctggttagtaatggtcagagatcaagat 3238  
Db 3096 ATCAATGGCTATGTATGATGATACACTCCCTGGCTTATATGCTCAGAAATCAAGATC 3155  
QY 3239 cgaatgtagtctgctcagatgagcagatgaaacatccatctattcattctcagtga 3298  
Db 3156 CGATGATATCTGCTCAGATGGGAGCAATGAAATATCCATTCGATTCATTTTATGCGGA 3215  
QY 3299 catgtgtcactgtcagaaagaaagagatataaatgycacttataatccatccca 3358  
Db 3216 CAGCTGTTCAGTATGCGAAAGAGAGAGATATAAATGGCCGTGTACATCTCTATCCG 3275  
QY 3359 ggtgtttttagagacatggaatgttaccatccaaagcttgaattggcgggtggaatgc 3418  
Db 3276 GGTGTCTTTGAGACAGTGAATGCTACCGTCAAGTTGGAATTTGGCGAATGAATGC 3335  
QY 3419 ctatagtgagagcatctacatctggagtagagcactttctgtgttaagagataag 3478  
Db 3336 CTGATTTGGCGACACTGTCAGAGCTGGATGAGACGACTTTCCTGTGTACGACGAGAG 3395  
QY 3479 tgltagaactccctggagaaatgctctgtagacatataagatlltcaagattacagctca 3538  
Db 3396 TGTACAGCTCAGTGGGAATGGCTTTGAGCAGATTAAGATTTTCAGATCACACTTCA 3455  
QY 3539 ggaacatatgagacagtgagcccaagctggccagacttcatatccggatcaatcaat 3598  
Db 3456 GGACAGATATGACAGTGGGCCCAAGCTGGCCACACTTCAATTAATCCGGATCAATCAT 3515  
QY 3599 gcccggagacacaaagagcccttcttggatcgaagtgagatctgttggcaccatgatt 3658  
Db 3516 GCCTGGAGACCAAGATCCCACTCTTGATCAAGTGTGATCTTTGGACCAATGATC 3575  
QY 3659 attacggcatcaagaccaaagtgcccgtagaagttctcgaagcctctacatctctag 3718  
Db 3576 ATTACGGCATCATAGCCCAAGGTGGCGGTGAGAAATTTTCAGAGCTCTGATCTGCCAG 3635  
QY 3719 ttatcatcatgtagtctttagtgggaagatgagcagacttalcgaggaaattccact 3778  
Db 3636 TTTATCATCATAGTACAGTCTTGACGGAGAACTGGCAGAGTTACCGAGGGAATTCACAG 3695  
QY 3779 ggaaccttaagtgcttcttggcaatgtgatatcatctggaataaacaataatcttt 3838  
Db 3696 GGCACTTAATGGTCTCTTTGGCAATGTGACGATCTGGGATTAACACATATATTTT 3755  
QY 3839 aacctccaattatctgcgataacatcgcttgcaccaactcatatagacatcgagc 3898  
Db 3756 AACCTCGAATGTGGCTGTACATCGTTTGACCCCAACACATTTACAGATCGCGAGC 3815  
QY 3899 actcttcgcatgagtgtagtggtctgtgattaaatagttgcagcagatgcatggaatg 3958  
Db 3816 ACTTTTGCATGAGAGTGTGATGGCTGATTTAAACATTTGACAGATGCCCTGGGAGATG 3875  
QY 3959 gagaagtaagcaataagatgagcaagattactgcttcatctacttaacaattgttt 4018  
Db 3876 CAGATTAAGGATATCAGATCTCAGATCAAGGCTCTCCACCTTAAGCAATATATTT 3935

QY 4019 gccacctgtgtctccttcaaaagctcgacttcaactccaaggagagtaatgcttgaga 4078  
Db 3936 GCCACTGTGTCCTCTTACAAAGCCGAGCTTCACTCCAGGGCGAGCAATGCCGTGGCGA 3995  
QY 4079 cctcagtggaataatccaaaagatggtgtgtagaagtgagtagtccgaagacaaatgaatc 4138  
Db 3996 CCCCggTACAGCGAGAGAGTGTGTGACAGTGTGACAGTGTGACAGTGTGACAGTGTGAC 4055  
QY 4139 acaagagtaactactcaaggagtaaaatctgtcttaccagcatgtagtgaagagtc 4198  
Db 4056 ACAGGCATCACCAACCGAGCGGTGAAGTCCCTGCTCAGACACATGTATGAAAGATTC 4115  
QY 4199 ctcatctccagcagtagcaagatgagcatagtggaactcttlttcagaatggaagta 4258  
Db 4116 CTGCTGTACAGTACAGAGAGCGCGCGCTGACACCTGTTCTTTCAGAGACGCGCACAGC 4175  
QY 4259 aaggttttcaagggaataatcaagactcttcaacactgtgtgtagactcttagaccagc 4318  
Db 4176 AAGGTTTTTCAAGGGCAATCAGAGACTCTCCACCCCGGTGTGAAGGCTGTGGACCCCGC 4235  
QY 4319 ttactgactcgtactcctcgaaattcaaaccccaagatgtgggtgacagagattgacctgag 4378  
Db 4236 CTGTTACAGCGCTTACTAGATCCACCCACGAGCTGGGCGCAGACATCGCCTTAGG 4295  
QY 4379 atgagagttctggtctgcgagcacaagagacttactga 4417  
Db 4296 CTCGAGGTTCTAAGATGTGAGGACAGAGATCTCTACATGA 4334

## RESULT 13

US-08-717-294-42  
: Sequence 42, Application US/08717294  
: Patent No. 6114148

## GENERAL INFORMATION:

APPLICANT: SEED, BRIAN  
APPLICANT: HAAS, JURGEN  
TITLE OF INVENTION: HIGH LEVEL EXPRESSION OF  
TITLE OF INVENTION: PROTEINS  
NUMBER OF SEQUENCES: 110  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Clark & Elbing LLP  
STREET: 176 Federal Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110

## COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/717,294  
FILING DATE: 20-SEP-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:

## ATTORNEY/AGENT INFORMATION:

NAME: Elbing, Karen L  
REGISTRATION NUMBER: 35,238  
REFERENCE/DOCKET NUMBER: 00786/345001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-428-0200  
TELEFAX: 617-428-7045  
TELEX:

## INFORMATION FOR SEQ ID NO: 42:

SEQUENCE CHARACTERISTICS:  
LENGTH: 4451 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear







Dd	2124	TCATGAGAGAACCCCGGCTGTGGATTCTTGGGGCTGCCAACAGACGACTTCCGACACCCG	218
Oy	2159	gscatgacogccttactgaaggcttcagtbtgbyacaagaacactgylgtaltatcag	2218
Dd	2184	GGCATGACTGCGCTGTGTAAGTCTCCAGCTGCCACAAGAACACCGCGGACTACTAGAG	2243
Oy	2219	gacagtataaagatrtatgaacatcttgctgtggtataaaacaaatgcattgaacccaaga	2278
Dd	2244	GACAGCTACGAGACATCTCCGCTACTCTGTGTCCAAAGAACACGCGCATGAGCCCGCG	2303
Oy	2279	agcttccccaagaattcaagacaacccctcagcacttgaagaaagcaattlaegyacaacca	2338
Dd	2304	TCTCTTCCCAAAACTCCGCCACACCAGCAGCGGTGAGAAGCATTCMAAGCCACCCCC	2363
Oy	2339	ccagctcttgaacgcacatcaacggygaataacttgctactcttcactcagtcagatcaag	2398
Dd	2364	CCCGTCTGMAAGCGCCACCGCGGAGATCACCCGCACACCTTGCAAGCGCACAGAG	2423
Oy	2399	gaattgacatgatgatataccataatcgtttgaaatgaaagaaagatttgacatttat	2458
Dd	2424	GGAATGACTACGACGACACCATTCAGCTGGAGATGAAAGAGGAGACTTGACATCTAC	2483
Oy	2459	galbyagatgaataatcagagccccgcgcagcttccaagaanaacagacatttla	2518
Dd	2484	GACGAGAGCAGAGAACGACGAGCCCGCGTCCCTCCAAAAGAAACCGCGCACTTACTC	2543
Oy	2519	gctgcagtggagaagctctggtgattatgtaatgtagtagtctcccaatgltctaaagaac	2578
Dd	2544	GGCGCGGTGAGCGCTGTGGGACTACGGCATGACGACGACCCCCACGTCCTCGCMA	2603
Oy	2579	agggcctcagtgagcagtgctccctcagttcaagaagttgtttccagaaattacgat	2638
Dd	2604	CGCGCCAGAGCGGACGCGTGCCTCCCACTTCAGAGAGTGTGTCCAGAGACTTCACGAC	2663
Oy	2639	ggctccttactcagcccttataccgtgygaagaaatgaacaattgygactcctg99g	2698
Dd	2664	GGAGCTTACCCAGCCGCTGTACCGCGGGAGCTGAACGAGCACTGGGCTGCTCGGC	2723
Oy	2699	ccaatataagagcagaagttgaagaataataatcgttlaacttcagaatatcagcctct	2758
Dd	2724	CCCTACATCGGCCCGGAGTGAGGACMACATCATGTGTGACTTCCGACCAAGCCCTCC	2783
Oy	2759	gctccattctcttactatctgaacttatcttatatagaagatcgaagcagaagca	2818
Dd	2784	CGGCTTACTCTTCTACTCTCTCCCTGATACGCTACGAGGAGGACAGCGCCAGGCGCC	2843
Oy	2819	gaactagaaaaaacttgcgaagcctlaaelyaaaccaaaacttacttltgaaagtgca	2878
Dd	2844	GAGCCCGCAGAACTTCGTGAAGCCCAAGAGACCTAAAGACTTCTTGGAAGTGCAG	2903
Oy	2879	calcatatgaccaccaataagatlgagtltgaactgcaagcctvggtctaltctctgat	2938
Dd	2904	CACCAATGCCCCCACCAGAGAGAGATTGTGACTGCAAGGCTCTGGGCTTACTTACGAGAC	2963
Oy	2939	gttgaccttgaaaaaaagtgcactcaagsgctatbtgagccctctgtgtctgcaact	2998
Dd	2964	GTGGACCTGGAGAAAGACGTGCACAGCGGCTTGATCGGCCCTTGCTGTGTGTGCACACC	3023
Oy	2999	aaacaactgaacctgtctcaatggygaacaaagtgaagatagacgaattgtctgttctc	3058
Dd	3024	AACACCTGAACCCCCCGGAGGCGAGGTGACTGTGCAGGAATTTGCCCTGTCTCTC	3083
Oy	3059	accactcttgaatgaagacaaagactgtgtacttcaactcyaaaaatargaagaagactgcag	3118
Dd	3084	ACCATTGTGACAGAGACTTAAGACTGTGTACTTCCACGAGAAACATGGAGCGCMATGCGCG	3143
Oy	3119	gctccctgaatbtccaagtatgaagatcccaactttaaagaaatbtgcttccatgca	3178
Dd	3144	GGCCCTTGCAACTCCAGATGGAAGATCCCACTTCAAGGAGAACTACCGCTTCCAGCGC	3203
Oy	3179	atcaatgctacaataatgtaataactcctggtcttaagttaatgctcagatcaagaagt	3238
Dd	3204	ATCAAGGCTACTCATMTGAGACACCTTCGCGCTGTGTATGTGCCCAAGACAGCAGCATC	3263

QY	3229	cgatggtatctggtccagcatgagggaagaatgaaacaatcatctatctatctaatggtga	3228
Db	3264	GCCTGGTACCTCGCTCTATTGTTGGGAGCAAGCAACATCCACGCTTCACCTTCAGCGGC	3323
QY	3299	catgtgtcactctgacgaaaaaaagagagatataaaatgagcactgtacatctcatcca	3358
Db	3324	CACGTTTTCACCGTGGCGCAAGAAAGAGAGATACAGATGATGGCCTGTACAACTGTACCC	3383
QY	3359	ggtgttttttgagacagtggaaatgttacccaaccaagctgtgaatttgcgtggtgaatgc	3418
Db	3384	GGCGGTTCGAGAGCTGTGGAGATCTGTCCGACGAAAGCGGGAGATCTGGGCGCTGGAGTGC	3443
QY	3419	cttatgtcgagacatacacaatgtcgtggagatagacacacttttctgtgttaagaaataat	3478
Db	3444	CTGATCGCGGAGCAACTGTGCACGCGCGGATAGCACTCTGTCTGTGTACAGCAACAG	3503
QY	3479	tgtaagactccctctgggaatgtgtctctgtgacacatlagagatttcaagatcaagctca	3538
Db	3504	TGCGCAGACCCCTCGGGCATATGGCCAGCGGCCCATCTCGGACTTCAGATCACCCGACGC	3563
QY	3539	ggacaataatgagcaggtggccccaagctctgcagcagctcatattctcgatcaatacat	3588
Db	3564	GGCCAGTACGGCCAGTGGGCTCCCAAGCTGGCCGCTGACTACAGCGGAGCATCAAC	3623
QY	3599	gccttggagcaacaagagaccccttctcttgatcaaaagtgtgactctgtgcacaatgatt	3658
Db	3624	GCCTGTGTGACCAAGAGAGCCCTTCTCTGATCAAGTGTGAGCATCTGTGGCCCATGATC	3683
QY	3659	altcaagcatcaagaccaccaaggtgtcccgatcaaaatctccaagctctacatctctatg	3718
Db	3684	ATCCACGGCATCAAGACCCAGGGGGCCCGCCACAAATTTAGAGACCTGTACATTCAGCCAG	3743
QY	3719	tttaatcaatcgtatagtctcttgatggaaagatggcagactatccgaagaattccact	3778
Db	3744	TTTCAATCATCATGTACTCTCTTAGAAGGCAAAATGTGGCAACCTACCGGGGCAACGACC	3803
QY	3779	ggaacctaaatggtctctcttcttgccaatgttgatcaatctgtgataaacaacataatttt	3838
Db	3804	GGCACCCCTGATGGTGTCTTTCGGCAACGTGGACAGCAGCGGATCAACACCAACATCTTC	3863
QY	3839	aacctccaatattgtctgaataaataacgtttgcaccaacatcatatagatttgcagc	3898
Db	3864	AACCCCCCATCATCTGCCCGCTCATCCGCTGCACCCACCCACATACGATCCGCGAGC	3923
QY	3899	actcttcgcatggagtgtgagctgtgatttaaatagttgacagatgcattggaaatg	3958
Db	3924	AOCCTGGCATGTGAGGTGATGGGTGGAGCTCAACAGCTGCAGCATGGCCCTGGGCATG	3983
QY	3959	gagagttaaagcaataatcaagatgcacagatctgcttcaactctacttaaccaatatgttt	4018
Db	3984	GAGAGCAAGGCGATCAACGACGCCCAATATACCGCTCAGGTACTTCAACCAATGTTC	4043
QY	4019	gcccactgtgtctcttcaaaagctgcagcttcaactccaagggaaggatgaatgctctggaga	4078
Db	4044	GCCACTGTGAGACCCCGACAGAGGCCGCGCTGCACCTGCAAGGCCGACGAAAGCCGTGGCC	4103
QY	4079	cctaagtgtataatccaaaaagtgtgctgcgaagtgtgacttccagaaagacaatgaagtc	4138
Db	4104	CCCCAGGTGAACAACCCCAAGAGATGGCTGTGCAGTGTGATCTTCAAAAAACATGAAAGTG	4163
QY	4139	acagagatgaacttaactgaaggagtaaaatctctgcttaccagagttaigtgaagagctc	4188
Db	4164	ACTGGCGTGACCAACCCAGGGCGTCAAGAGCCTGTCTACCAAGATGTACTGAAGAGTTTC	4223
QY	4199	ctcatctccacagatcaagaatggccatcagttgtgactctcttttccagaatgtgcaagta	4258
Db	4224	CTGATCAGACACACGACGAGCGGCAACAGTGAACCTGTCTTCCAAAAAGCGCAAGGTG	4283
QY	4259	aaggttttccaggaatatcaagactccttcaaacctgtgtgtgaactctctagaccacag	4318
Db	4284	AAGGTGTTCCAGGGCAACAGGAGACGTCTACACCGGTGTGTAAACAGCTGTGACCCCTCC	4343







Qy 1141 tgaatgcaaacctcctccttccatccaaatcgcgtcagctgcgaagaacatccctaaac 1200  
1290 tgaatgcaaacctcctccttccatccaaatcgcgtcagctgcgaagaacatccctaaac 1349  
Qy 1201 ttggatcatcatcctgctgcgtgaagaaggaacatgagcctatgctcccttagtctgcg 1260  
1350 ttggatcatcatcctgctgcgtgaagaaggaacatgagcctatgctcccttagtctgcg 1409  
Qy 1261 ccccgatgacagaagttataaagtcgaatctgaacaaatggccctcagcgagctggtag 1320  
1410 ccccgatgacagaagttataaagtcgaatctgaacaaatggccctcagcgagctggtag 1469  
Qy 1321 gaagtaacaaaagtcgcgtatctgcatcacagaatgaacatctgaagctcgtgaagc 1380  
1470 gaagtaacaaaagtcgcgtatctgcatcacagaatgaacatctgaagctcgtgaagc 1529  
Qy 1381 tattcagaatgaatgaagaatcttggaaccttactcttattggggaagcttggaacact 1440  
1530 tattcagaatgaatgaagaatcttggaaccttactcttattggggaagcttggaacact 1589  
Qy 1441 gttgattatattaagaatcaagaacagacacataatacctaccctcaggaatcac 1500  
1590 gttgattatattaagaatcaagaacagacacataatacctaccctcaggaatcac 1649  
Qy 1501 tgaatgcccgcctctgtatctcaaggagatcaacaaaaggctgaacatctgaagatct 1560  
1650 tgaatgcccgcctctgtatctcaaggagatcaacaaaaggctgaacatctgaagatct 1709  
Qy 1561 tccaattctgcgaagagaataatcaataataatgaagaagctgctgaagaatggcc 1620  
1710 tccaattctgcgaagagaataatcaataataatgaagaagctgctgaagaatggcc 1769  
Qy 1621 aacaaatcagatcctgcgtgcctcagccgctactactagcttctgtaataatgaag 1680  
1770 aacaaatcagatcctgcgtgcctcagccgctactactagcttctgtaataatgaag 1829  
Qy 1681 agatctagctcaggaactatgcccctcctcatalctgctacaaagaatctgtagatca 1740  
1830 agatctagctcaggaactatgcccctcctcatalctgctacaaagaatctgtagatca 1889  
Qy 1741 aagaggaaacagataaagtcaagaagaagaatgcatcctgcttctgtaattgata 1800  
1890 aagaggaaacagataaagtcaagaagaagaatgcatcctgcttctgtaattgata 1949  
Qy 1801 gaacgaagctggtatcccaacagagaataatacaagcttctccccaacgaagctgagt 1860  
1950 gaacgaagctggtatcccaacagagaataatacaagcttctccccaacgaagctgagt 2009  
Qy 1861 gcaagcttgaagatccagaagcttccaaagctcacaatcatgcaagcatcaatgctatgt 1920  
2010 gcaagcttgaagatccagaagcttccaaagctcacaatcatgcaagcatcaatgctatgt 2069  
Qy 1921 ttgtgtagtctgcaagctgctgctgcttctgtaagaagtgcgaatctgtaacatctaa 1980  
2070 ttgtgtagtctgcaagctgctgctgcttctgtaagaagtgcgaatctgtaacatctaa 2129  
Qy 1981 catcgaagcacagaagctcctcctcctgcttctcctcctgtaatacttaacaaacaa 2040  
2130 catcgaagcacagaagctcctcctcctgcttctcctcctgtaatacttaacaaacaa 2189  
Qy 2041 aatgctctatgaagaacacactcaactatcccatctcgaagaagaactgctctatgtc 2100  
2190 aatgctctatgaagaacacactcaactatcccatctcgaagaagaactgctctatgtc 2249  
Qy 2101 gatggaaaaacccaggtctcatggaatctcggtgcgaacactagactctcggaacagag 2160  
2250 gatggaaaaacccaggtctcatggaatctcggtgcgaacactagactctcggaacagag 2309  
Qy 2161 catgacccgcttactgaaggtcttctgctgctgtaagaagaacactggtatatacagaag 2220  
2310 catgacccgcttactgaaggtcttctgctgctgtaagaagaacactggtatatacagaag 2369

Qy 2221 cagttatgaagatattcagcatctgctgagtaaaaaaagcattgaaccagaag 2280  
Db 2370 cagttatgaagatattcagcatctgctgagtaaaaaaagcattgaaccagaag 2429  
Qy 2281 ctctcccaagattcaagcaccctagcctagggcaaaagcattatgcccacca 2338  
Db 2430 ctctcccaagattcaagcaccctagcctagggcaaaagcattatgcccacca 2487

RESULT 15  
US-08-251-937A-3  
; Sequence 3, Application US/08251937A  
; Patent No. 5583209  
; GENERAL INFORMATION:  
; APPLICANT: Lollar, John S.  
; APPLICANT: Runge, Marschall S.  
; TITLE OF INVENTION: Hybrid Human/Porcine Factor VIII  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Kilpatrick & Cody  
; STREET: 1100 Peachtree Street  
; CITY: Atlanta  
; STATE: Georgia  
; COUNTRY: US  
; ZIP: 30309  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/251,937A  
; FILING DATE: 31-MAY-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/864,004  
; FILING DATE: 07-APR-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Pratl, John S.  
; REGISTRATION NUMBER: 29,476  
; REFERENCE/DOCKET NUMBER: EMU106DIY  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 404-815-6567  
; TELEFAX: 404-815-6555  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9009 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: N-terminal  
; ORIGINAL SOURCE:  
; ORGANISM: Homo sapien  
; TISSUE TYPE: Liver  
; FEATURE:  
; NAME/KEY: misc-feature (Domain Structure)  
; LOCATION: 5001...7053  
; OTHER INFORMATION: /note= "Equivalent to the A3-C1-C2  
; OTHER INFORMATION: domain"  
; FEATURE:  
; NAME/KEY: misc-feature (Domain Structure)  
; LOCATION: 1...2277  
; OTHER INFORMATION: /note= "Equivalent to the A1-A2  
; OTHER INFORMATION: domain"  
; US-08-251-937A-3

Query Match 52.9%; Score 2336.4; DB 1; Length 9009;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2337; Conservative 0; Mismatches 1; Indels 0; Gaps 0;



QY	1	catgcaaatagagctctccacgtgcttctgtgaccttttgagcttgcgattgcttagtgc	60
DB	150	CATGCAAAATAGAGCTCTCCACCTGCTCTTTCTGTGCTTTTGCCATTCGTCTTAGTGC	209
QY	61	caaccagaatactacccctggtgtgcagtgtgaactgtcatatgagactataagcaagtatct	120
DB	210	CACCAAGAATACCTACCTGGGTGCGATGGAACGTATGATGGAGCTATATGCAAAAGATCT	269
QY	121	cgggtgagctgctgtgtgagcgaagattccctctagagatgtgccaaaactcttccatcaa	180
DB	270	CGGTAGCTGCTGTGGAGCGCAAGATTTCTCTAGAGTGCCTCAAAATCTTTTCCATTCAA	329
QY	181	cacccagctcgtgtacaataaagacatctgtttgtaaatccaggttcaaccttttcaaat	240
DB	330	CACCTCATGCTGCTACAAAAGACTCTGTTGTAGAAATTCACGGTTTCACCTTTTAACT	389
QY	241	cgttaagccaagcgcacccctgtgatgtgtcgaagtcctcaaccacccaagcgtgaagtta	300
DB	390	CGCTAAGCCAAGGCCACCTGGATGGTCTGCTAGTCTTACCATTCAGGCTGAGGTTTA	449
QY	301	tgtatacagtgatcatataacttaagaacatggtctccatccctgtcagttctcatgctgt	360
DB	450	TGATACAGTGTGCTATACCTTAAGACATGGCTCCCATCTGTCAGTCTTTCATGCTGT	509
QY	361	tgtgtatcctactgtgaagacttctgaaggagctgaatataatgataagaccagtcgaag	420
DB	510	TGTGTATCTTACTTCTGGAAGCTTCTGTAGGAGACTAATATGATGATGACACAGTCAAG	569
QY	421	ggaagaagaagaatgaataagcttccctgtgtggaagccatacatagtctcgtgagtgct	480
DB	570	GGAGAAAGAAGATGATTAAGTCTTCCCTGTGGAGAGCCATATGATGTCTGGACAGTCT	629
QY	481	gaagaagaatgttccaatgtgacctgtgaaccacatgtgtccttactactatattcttcca	540
DB	630	GAAAGAGATGCTCCAAATGGCTCTGACCCACTGCTTACTACTATATCTTTCTCA	689
QY	541	tgtgaacctgtgtaaaagacttgaattcagacctcatltggaagccctactagatgtagaga	600
DB	690	TGTGACCTGTGTAAGAAAGCTTGAATTCAGGCTCATTTGGAGCCCTACATGATGAGGA	749
QY	601	aggaagcttgcgccaaggaagaacacagaccttgcacaattatactacttctgtcgt	660
DB	750	AGGGGCTGTGGCCAAAGGAACACACAGACCTTGACAAATTTATACATTTTGTCTGT	809
QY	661	attgataagaaggaaagtgtgcacctcagaacaaagaacctcttgatgcagagataggga	720
DB	810	ATTGATGAAAGGAAAGTTGGCACTCAGAAACAAAGAACTCTTGATGACAGATAGGGA	869
QY	721	tgtcgcacatcgtcgtgagcctgtgcttaaaatgcacacagtcaatggttattgtaaacagtc	780
DB	870	TGCTGCAATCTCTCGGGCTGGCTTAATAATGCACACAGTCAATGGTTATGTAACAGTCT	929
QY	781	tctgtccaggttcatgtgatgtgccacaaggaatacagttcatatgtgacatgtgatgtgag	840
DB	930	TCTCCAGGCTCTGATTTGATGCCACAGGAATTCAGTATTTGCAATGATTTGGAATGG	989
QY	841	caccacccctgaagtgtacataataatctcctgaaggtacacatcttctgtgaggaacca	900
DB	990	CACACCTCTGTAAGTGCACCTAATATTCCTCGAAGGTCAACATTTCTTGTGAGGAACCA	1049
QY	901	tgcgcaagcgctccttggaaatctgcgcaataacttctcttactgtcctcaaaccttgat	960
DB	1050	TGCGCAGGCGCTTGTGGAATCTGCGCAATATCTTCTTACTGCTCAAAACCTTTGAT	1109
QY	961	ggaaccttggacagttctactgttttgcatatactcttccacaacaatgaatgtgcatgga	1020
DB	1110	GGACCTTGGACAGTTTCTACTGTTTGTGATATCTTCTCCACCAACATGATGACATGGA	1169
QY	1021	agctatgtcaaatagacagctgtccagaggaagcccaactaagaatgaataaatatga	1080
DB	1170	AGCTTATGTCAAAGTACAGCTGTCCAGAGAACCCCACTACGAATGAAATAATATGA	1239
QY	1081	agaagcgaagaactatgatgatcatcttacttccctgaagaatggtgtgcagtttga	1140
DB	1230	AGAAAGCGAAGACTATGATGATGATCTTACTGATTTCTGAATGATGATGTCAGGTTTGA	1289
QY	1141	tgatgacaacttcccttcttccataccaatcgtctcagttgtgccaagaacatccctaaac	1200
DB	1290	TGATGACAACCTCTCCCTCTTATATCAAAATTCGCTCAGTTGGCAGAAACATCTTAAAC	1349
QY	1201	tgtgtgatactacatctgtcgtgaagaggaagacttggagactatgtcccttagtccctgc	1260
DB	1350	TTGGGTACATTTACATTTGCTGCTGAAGAGGAGACTGGGACTATGCTCCCTTATGCTCCG	1409
QY	1261	cccagatcagaagaattataaagaatatttgaacaaatggccctcagcgagttgtgag	1320
DB	1410	CCCCGATGACAGAGTTATTAAGTCAATATTTGAACATTTGSCCTCACCGGATGTAG	1469
QY	1321	gaagtaacaaaagtcocgaattatgtgcatacagaatgaagaaccttaagaacctgtgaagc	1380
DB	1470	GAAGTACAAAAAAGTCCGATTTATGGCATACACAGATGAACCTTTAAGACTCGTGAAC	1529
QY	1381	tattcaagcatgaatgaagaatcttgggaccttacttatttggggaagttggaacaact	1440
DB	1530	TATTGAGATGATATCAGGAATCTTGGGACCTTACTTTATGGGGAAGTTGAGACACT	1589
QY	1441	gttgaattatattgaagaatcaagcagaagaccataataacatctacccctcaggaatac	1500
DB	1590	GTTGATTTATTTAAGAAATCAAGCAAGCAGACATATTAACATCTACCCCTACGGAAATC	1649
QY	1501	tgaatccgctccttgtatccaagagatltacaaaaggtgtgtaaaacatltgaagaatt	1560
DB	1650	TGATGCTCGTCTCTTGTATTCAGAGAGATTTACCAAAAGGTGTAAACATTTAAGATTT	1709
QY	1561	tccaattctgcaggaagaataatccaataataatgtgacagtgactgtgaagaatgtggcc	1620
DB	1710	TCCAATTTCTGCAGAGAAATATTCAAATATTAAGAGACAGTGTGAAGAATGGGCT	1769
QY	1621	aactaatcagaatccctgtgtgtgcctcagccgctattaccctgaatttcgttaataatgaag	1680
DB	1770	AACTAATTCAGATCTCTGGTGCCTAACCCGCTATTTACTCTATTTCTGTTAAATGAGAG	1829
QY	1681	agatctagcttccaagatcattatgtgcctctctcaatcgtcctacaaagaatcgtatagca	1740
DB	1830	AGATCTACTCTCAGAGATCATTTGGCCCTCTCTCATGTGCTACAAAGATGCTGATATCA	1889
QY	1741	aagaggaacacagataaagtgtcaagaagaagaatgtcatcctgtttctgtatcttgatga	1800
DB	1890	AAGAGGAACCAAGATTAATGTACAGACAAAGGAATGTCAATCTGTTTCTGTTATTTGATGA	1949
QY	1801	gaaccgaagctgtgtaccccaagaagaataatacagcttctccccaatccagctggagct	1860
DB	1950	GAACCGAAGCTGTGTAATCTCACAGAGAAATATACAGCTTTCTCCCAATCCAGCTGAGT	2009
QY	1861	gcagcttgaagatccagaagttcccaagcctccaacatcaatgtcagatgtcatgt	1920
DB	2010	GCAGCTTGAAGATCCAGAGTTCCAAAGCCTCCAAATCATGACACATCAATGGCTATGCT	2069
QY	1921	tttggatagtttgcaggtgtgcaagttgttcttgcagaggtgtgcatatctgataag	1980
DB	2070	TTTTGATATGTTTGCAGTTGTGATTTTGTTCATGAGTGGCATCTGTCATATTTCTAAG	2129
QY	1981	cattggaagcacagacttgaacttcttctgttcttctcctgtgatatctactccaacacaa	2040
DB	2130	CATTGGACACAGACTGACTTCTTCTTCTGTTCTTCTCTGATATATCTTCAAAACACAA	2189
QY	2041	aatgtctatgaagaacacactcaaccttatccatcttccaaggaagaactgtcttcatgtc	2100
DB	2190	AATGCTTATGAAGACACACTACCCCTATTCCTCAATTCAGAGAAACTGTCTTATGATCT	2249
QY	2101	gatggaaaacccaggtctatgtatcttggggtgtgccaaactcgaacttgcgaagcagagg	2160
DB	2250	GATGGAAAACCCAGGTCTATGATTTCTGGGGTGCACAACTCAAGCTTTCGGAACAGAGG	2309
QY	2161	catgaccgaccttactgaaggttctcagttgtgataagaacacbtgtgtattatccagaga	2220



Db	2310	CATTGACCGCCTTACTGTAAGTTTCTAGTTGTGTGACAGSAAACACTGGTGATTTATTATTCGAGGA	2369
Oy	2221	cagttatgaaagatatttcagcatactctgctgagtaaaacaatgccaattgaaccaagaag	2280
Db	2370	CAGTTATGAAAGATTTTTCAGCACTACTGTCGAGATTAATAAACAAATGCCATTGAAACCAAAAG	2429
Oy	2281	ctctccccaagaattcaagacacccctagcacttagcacaagaacatttaatgcaccca	2338
Db	2430	CTTCTGCCGAATTCAGACACCCCTGACACTATAGCAAAACCAATTTAATGCGACACACA	2487

Search completed: November 18, 2001, 04:58:30  
Job time: 35512 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 18, 2001, 05:05:42 ; Search time 919.72 Seconds

(without alignments)  
3015.528 Million cell updates/sec

Title: US-09-689-430-1\_COPY\_419\_4835

Perfect score: 4417

Sequence: 1 catgcaatagagctctcca.....agccacagagcctctactga 4417

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 730101 seqs, 313950809 residues 1460202

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : N\_Geneseq\_0601.\*

- 1: /SIDS8/gcgdata/geneseq/geneseqn/NA1980.DAT.\*
- 2: /SIDS8/gcgdata/geneseq/geneseqn/NA1981.DAT.\*
- 3: /SIDS8/gcgdata/geneseq/geneseqn/NA1982.DAT.\*
- 4: /SIDS8/gcgdata/geneseq/geneseqn/NA1983.DAT.\*
- 5: /SIDS8/gcgdata/geneseq/geneseqn/NA1984.DAT.\*
- 6: /SIDS8/gcgdata/geneseq/geneseqn/NA1985.DAT.\*
- 7: /SIDS8/gcgdata/geneseq/geneseqn/NA1986.DAT.\*
- 8: /SIDS8/gcgdata/geneseq/geneseqn/NA1987.DAT.\*
- 9: /SIDS8/gcgdata/geneseq/geneseqn/NA1988.DAT.\*
- 10: /SIDS8/gcgdata/geneseq/geneseqn/NA1989.DAT.\*
- 11: /SIDS8/gcgdata/geneseq/geneseqn/NA1990.DAT.\*
- 12: /SIDS8/gcgdata/geneseq/geneseqn/NA1991.DAT.\*
- 13: /SIDS8/gcgdata/geneseq/geneseqn/NA1992.DAT.\*
- 14: /SIDS8/gcgdata/geneseq/geneseqn/NA1993.DAT.\*
- 15: /SIDS8/gcgdata/geneseq/geneseqn/NA1994.DAT.\*
- 16: /SIDS8/gcgdata/geneseq/geneseqn/NA1995.DAT.\*
- 17: /SIDS8/gcgdata/geneseq/geneseqn/NA1996.DAT.\*
- 18: /SIDS8/gcgdata/geneseq/geneseqn/NA1997.DAT.\*
- 19: /SIDS8/gcgdata/geneseq/geneseqn/NA1998.DAT.\*
- 20: /SIDS8/gcgdata/geneseq/geneseqn/NA1999.DAT.\*
- 21: /SIDS8/gcgdata/geneseq/geneseqn/NA2000.DAT.\*
- 22: /SIDS8/gcgdata/geneseq/geneseqn/NA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	4412.8	99.9	4629	20	AAx88293	Human Factor VIII
2	4407.4	99.8	4629	18	AAx73164	CDNA encoding huma
3	4355.2	98.6	4629	16	AAO76016	B-domain deleted F
4	4351	98.5	4670	19	AAV23339	Human Factor VIII
5	4321.4	97.8	4832	19	AAV19581	Human Factor VIII
6	4321.4	97.8	4832	19	AAV15338	Human Factor VIII
7	4321.4	97.8	4832	19	AAV15338	Human Factor VIII
8	4319.8	97.7	4999	21	AAx82259	Beta-domain delete
9	4315	97.7	1245	21	AAx82259	Recombinant adeno
10	4249.4	96.2	11933	21	AAx82259	Vector H5QReNeo fo
11	4173	94.5	4275	10	AAx90654	Recombinant adeno
						DNA encoding 740 A

12	4171.6	94.4	4275	9	AAx80446	Modified factor VI
13	4170	94.4	4272	9	AAx80447	Modified factor VI
14	4161	94.2	4373	20	AAx82258	Beta-domain delete
15	4152	94.0	4830	9	AAx81544	Human Factor VIII-
16	4108.8	93.0	4545	9	AAx80444	Modified factor VI
17	3743.8	84.8	5035	18	AAx69811	Factor VIII-dB695-
18	3708.4	84.0	4616	9	AAx81545	Human Factor VIII-
19	3678	83.3	5094	21	AAx49231	Human Factor VIII-
20	3189.8	72.2	4334	19	AAV12142	DNA construct H5Q/
21	3189.8	72.2	4334	20	AAx91195	Homo sapiens facto
22	3189.8	72.2	4334	22	AAx91195	Porcine factor VII
23	2603.8	58.9	4451	19	AAx90541	CDNA encoding porc
24	2336.4	52.9	9009	14	AAV23288	Synthetic human Fa
25	2336.4	52.9	9009	18	AAO50185	Human Factor VIII
26	2336.4	52.9	9009	19	AAV25810	Factor VIII:C (Arg
27	2336.4	52.9	9009	20	AAx91162	Human factor VIII
28	2336.4	52.9	9009	22	AAx90508	Human factor VIII
29	2334.8	52.9	6300	17	AAx703571	Factor VIII CDNA
30	2334.8	52.9	8241	9	AAx81439	Factor VIII CDNA
31	2334.8	52.9	8241	9	AAx81096	Factor VIII CDNA
32	2334.8	52.9	8241	9	AAx81096	Factor VIII CDNA
33	2334.8	52.9	8975	6	AAx50054	Human factor VIII
34	2334.8	52.9	8975	21	AAx38604	Human full-length
35	2334.8	52.9	9029	22	AAx60309	Human factor VIII
36	2334.8	52.9	9068	19	AAV15359	Human factor VIII
37	2334.8	52.9	9080	19	AAV19580	Human factor III e
38	2333.8	52.9	11846	20	AAx82261	Factor VIII protel
39	2333.2	52.8	7053	18	AAx51357	Factor VIII:C code
40	2333.2	52.8	8967	17	AAx31031	Factor VIII full-l
41	2333.2	52.8	8967	22	AAx87526	Factor VIII full-l
42	2332.2	52.8	9009	19	AAV18884	Human factor VIII
43	2332.2	52.6	7056	15	AAO6615	Sequence of human
44	2309.8	52.3	7440	7	AAx60689	Sequence encoding
45	2272.4	51.4	7440	20	AAx82260	Factor VIII protel
						DNA sequence enco

## ALIGNMENTS

RESULT	1	
AAx88293	standard; DNA: 4629 BP.	
XX		
AC	AAx88293:	
XX		
DT	24-SEP-1999 (first entry)	
XX		
DE	Human Factor VIII with B-domain deleted CDNA.	
XX		
KW	Adenoviral vector: Factor VIII; Factor IX; clotting factor; treatment;	
KW	haemostatic; hemophilia A; haemophilia B; gene therapy; ss.	
OS	Homo sapiens.	
XX		
PN	US5935935-A.	
XX		
PD	10-AUG-1999.	
XX		
PF	07-JUN-1995; 9505-0484891.	
XX		
PR	07-JUN-1995; 9505-0484891.	
PR	10-JUN-1993; 93US-0074920.	
PR	25-MAR-1994; 94US-0218335.	
PA	(GENE-) GENETIC THERAPY INC.	
PI	Connolly S, Kalexo M, Smith T;	
XX		
XX	WPI; 1999-457617/38.	
DR		
XX		
PT	Adenoviral vectors useful for treating hemophilia	
XX		
PS	Example 1; Column 53-58; 90pp; English.	



XX This invention describes novel adenoviral vectors comprising at least  
CC one DNA sequence encoding a clotting factor (Factor IX or Factor VIII).  
CC The vectors of the invention have haemostatic activity. The vectors are  
CC useful for the treatment of hemophilia A or hemophilia B by gene therapy.  
CC This sequence represents human Factor VIII cDNA which has the B domain  
CC deleted.  
XX

SQ Sequence 4629 BP; 1319 A; 1050 C; 1015 G; 1245 T; 0 other;

Query Match 99.9%; Score 4412.8; DB 20; Length 4629;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 4414; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 atgcaaaatagagctctccacgtctctcttctgtgctcttttgagttcgattcgcttagtgc 61  
DB 1 atgcanaaagagctccaccacgtctctctctgtgctcttctgagattcgtcttagtgc 60  
OY 62 accagaagatactactctgtgacgtgagacgtgcatgtgagactatgcaagtgatctc 121  
DB 61 acaagaagatactactctgtgacgtgagacgtgcatgtgagactatgcaagtgatctc 120  
OY 122 ggtgagctgctgtgagacgaagattccctcctagagtgccaaatctttccattcaac 181  
DB 121 ggtgagctgctgtgagacgaagattccctcctagagtgccaaatctttccattcaac 180  
OY 182 acccagctcggtacaaaagaagctctgtttgttagattcaggttccacgttttcaac 241  
DB 181 acccagctcggtacaaaagaagctctgtttgttagattcaggttccacgttttcaac 240  
OY 242 gctaagcaagcgcaacctgtgagtgctgtctaggttccatccatccaggctgagttat 301  
DB 241 gctaagcaagcgcaacctgtgagtgctgtctaggttccatccatccaggctgagttat 300  
OY 302 gatacaatgtgcatcatcacttaagaacatgtgcttcccatctctgtcaagcttcatgtt 361  
DB 301 gatacaatgtgcatcatcacttaagaacatgtgcttcccatctctgtcaagcttcatgtt 360  
OY 362 ggttaccctactctgaaagctctctgaggagctgataatgatatcagacagccagcaagg 421  
DB 361 ggtgtaccctactctgaaagctctctgaggagctgataatgatatcagacagccagcaagg 420  
OY 422 gagaaagaaga tga taaagcttccctgtgtggaagcacaataatgcttgcaggtcctg 481  
DB 421 gagaaagaaga tga taaagcttccctgtgtggaagcacaataatgcttgcaggtcctg 480  
OY 482 aaagagaatgttccaatgagctctgacccactgtgcttactactatcatcttctcat 541  
DB 481 aaagagaatgttccaatgagctctgacccactgtgcttactactatcatcttctcat 540  
OY 542 gtgacactgtgcaaaaagacttaattcaagcctcatcttgagccctactagatgtagaa 601  
DB 541 gtgacactgtgcaaaaagacttaattcaagcctcatcttgagccctactatgttagaa 600  
OY 602 gggagctctggcacaagaaagacacagacctgtgcacaattatatactacttttgcgtla 661  
DB 601 gggagctctggcacaagaaagacacagacctgtgcacaattatatactacttttgcgtla 660  
OY 662 ttgtatgaagggaagaatgttgcaactcagaacaaagaactcctgtatgcagataggat 721  
DB 661 ttgtatgaagggaagaatgttgcaactcagaacaaagaactcctgtatgcagataggat 720  
OY 722 gctgcactgtctggcgctggcctaaatgtcacacagtcacatgtgtatgtaaacggtct 781  
DB 721 gctgcactgtctggcgctggcctaaatgtcacacagtcacatgtgtatgtaaacggtct 780  
OY 782 ctgcagagctctgattgtgatgcacaggaatcaagtctatctggcagtgatgtggaatgggc 841  
DB 781 ctgcagagctctgattgtgatgcacaggaatcaagtctatctggcagtgatgtggaatgggc 840  
OY 842 accactctgaagtgacactcaatatctctgaaagtgacacatcttctgtgaggaacat 901  
DB 842 accactctgaagtgacactcaatatctctgaaagtgacacatcttctgtgaggaacat 901

DB 841 accactctgaagtgacactcaatatctctgaaagtgacacatcttctgtgaggaacat 900  
OY 902 cgcagagctctcttggaatctcgccaataacttctactgtgtccaacacttgatg 961  
DB 901 cgcagagctctcttggaatctcgccaataacttctactgtgtccaacacttgatg 960  
OY 962 gaccttgcaggtctctactgtttgttcatactcttccccaacatgagtgcagaa 1021  
DB 961 gaccttgcaggtctctactgtttgttcatactcttccccaacatgagtgcagaa 1020  
OY 1022 gcttatgtcacaagtacagctgtgtccagaagaccccaactacgaatgaanaataatgaa 1081  
DB 1021 gcttatgtcacaagtacagctgtgtccagaagaccccaactacgaatgaanaataatgaa 1080  
OY 1082 gaagcggagagactatgatactgattctatcttgaaatggatgtgtgagttgat 1141  
DB 1081 gaagcggagagactatgatactgattctatcttgaaatggatgtgtgagttgat 1140  
OY 1142 gatgcaactctctctcttataccaaatgcgtcagtttgccaagaagcaltctaaact 1201  
DB 1141 gatgcaactctctctcttataccaaatgcgtcagtttgccaagaagcaltctaaact 1200  
OY 1202 tgggtacattacatgtgcgtcgaagagagagacgtggaactatgtctccctagctcgc 1261  
DB 1201 tgggtacattacatgtgcgtcgaagagagagacgtggaactatgtctccctagctcgc 1260  
OY 1262 cccgagtgacagaagctataaagaatcaatgttgacaatgctccacagcgattgtag 1321  
DB 1261 cccgagtgacagaagctataaagaatcaatgttgacaatgctccacagcgattgtag 1320  
OY 1322 aagtaaaaaaagtcogattatgtacacacagatgaacactttaaagctgtgaaact 1381  
DB 1321 aagtaaaaaaagtcogattatgtacacacagatgaacactttaaagctgtgaaact 1380  
OY 1382 attcagcattagacaggaatttggagcttacttattatggggaagtgtgagacaactg 1441  
DB 1381 attcagcattagacaggaatttggagcttacttattatggggaagtgtgagacaactg 1440  
OY 1442 ttgattatatatgaagaacaaagcagacacataatacactccctccogataact 1501  
DB 1441 ttgattatatatgaagaacaaagcagacacataatacactccctccogataact 1500  
OY 1502 gatgtccgtccttgtatccaagagatcccaaaagtgtaaaacatttgaagattt 1561  
DB 1501 gatgtccgtccttgtatccaagagatcccaaaagtgtaaaacatttgaagattt 1560  
OY 1562 ccaattctgcagaggaatattcaaatatgaatggagaagatgtgagagggcca 1621  
DB 1561 ccaattctgcagaggaatattcaaatatgaatggagaagatgtgagagggcca 1620  
OY 1622 actaaatcagatccctcgggtgcctgacccgctattacttaatttcgttaatatgaaaga 1681  
DB 1621 actaaatcagatccctcgggtgcctgacccgctattacttaatttcgttaatatgaaaga 1680  
OY 1682 gatctagcttcaagatcatgtgcccctcctcatctgtctacaagaatctgtatgataa 1741  
DB 1681 gatctagcttcaagatcatgtgcccctcctcatctgtctacaagaatctgtatgataa 1740  
OY 1742 agaggaacacagataatgtacagacaagaggaatgcatccctggtttctgtatttgatgag 1801  
DB 1741 agaggaacacagataatgtacagacaagaggaatgcatccctggtttctgtatttgatgag 1800  
OY 1802 aaccgaagctgtactacctacagagaatataaagccttccccaatccagctgagtg 1861  
DB 1801 aaccgaagctgtactacctacagagaatataaagccttccccaatccagctgagtg 1860  
OY 1862 cagcttgcagatccagagttccaaagctccaaatcatgacagacatcaatgtctatgtc 1921  
DB 1861 cagcttgcagatccagagttccaaagctccaaatcatgacagacatcaatgtctatgtc 1920  
OY 1922 tttagatgttcaggttgcagttgtttgtgataagttgagatgagctgtaattcctaagc 1981  
DB 1921 tttagatgttcaggttgcagttgtttgtgataagttgagatgagctgtaattcctaagc 1980



[illegible]

QY	3062	atcttttgatgagccaaaagctggttacttcactcgaaataatggaaagaacatgcagggct	3121
Db	3061	atcttttgaatgagaccaaaagctgtgtacttcaactcgtaaaataatgaaagaacatgcagggct	3120
QY	3122	ccctgcgaatataccgaatgaggagatcccaacttttaagaagaattatcgcttccatgcaatc	3181
Db	3121	ccctgcgaatataccgaatgaggagatcccaacttttaagaagaattatcgcttccatgcaatc	3180
QY	3182	aatgctcaataatcggatatacctaccccgcttgatgaatggtcgaagatccaagaaggtatcga	3241
Db	3181	aatgctcaataatcggatatacctaccccgcttgatgaatggtcgaagatccaagaaggtatcga	3240
QY	3242	tgttatctgctcgaacatggtgcgaagaatgaaacacccatcttatcattcaattcagtyggacat	3301
Db	3241	tgttatctgctcgaacatggtgcgaagaatgaaacacccatcttatcattcaattcagtyggacat	3300
QY	3302	gtgttcacatgtgcaaaaaaaggagagtataaaatggcgcgttaacatgtctatcccaagt	3361
Db	3301	gtgttcacatgtgcaaaaaaaggagagtataaaatggcgcgttaacatgtctatcccaaccgct	3360
QY	3362	gttttttgacaagatggaataatgtatccctccaagaagcgggaatttgcgggttggaatgacct	3421
Db	3361	gttttttgacaagatggaataatgtatccctccaagaagcgggaatttgcgggttggaatgacct	3420
QY	3422	attgvcgaacatcacatctgcttgcgaatgtagcaacatttctcgtgtgtaacagaacaatagtgt	3481
Db	3421	attgvcgaacatcacatctgcttgcgaatgtagcaacatttctcgtgtgtaacagaacaatagtgt	3480
QY	3482	cagactcccccgggaatggtcttcgcgaacaattagagatttccagattccagcttcaga	3541
Db	3481	cagactcccccgggaatggtcttcgcgcacatataagatttccagattccagcttcaga	3540
QY	3542	caataatgacagatgggcccacaagaatgcgcagaactcatattccggatccaataatgac	3601
Db	3541	caataatgacagatgggcccacaagaatgcgcagaactcatattccggatccaataatgac	3600
QY	3602	tggagcacacaagaagacccttcttcgtgatccaagtggatctgttgcaccaaatgtatt	3661
Db	3601	tggagcacacaagaagacccttcttcgtgatccaagtggatctgttgcaccaaatgtatt	3660
QY	3662	cacggcatcaagaagaccgaaggtgcgcgttcgaagattctccaagctctacatctctagctt	3721
Db	3661	cacggcatcaagaagaccgaaggtgcgcgttcgaagattctccaagctctacatctctagctt	3720
QY	3722	atcatcatgtaatagtcttcttgaatggagaagatggtgcagactatacggagaatccaactgga	3781
Db	3721	atcatcatgtaatagtcttcttgaatggagaagatggtgcagactatacggagaatccaactgga	3780
QY	3782	accttaaatggtctctcttcttgcgaatgttgatatactctggataataaacacataattttaac	3841
Db	3781	accttaaatggtctctcttcttgcgaatgttgatatactctggataataaacacataattttaac	3840
QY	3842	cttccaattatgtgtcgaatacacatccgtttgcacccaacatataagatcttgcagacct	3901
Db	3841	cttccaattatgtgtcgaatacacatccgtttgcacccaacatataagatcttgcagacct	3900
QY	3902	cttcgcacatgagatggaatgaggtcgttgatgaataatagttgagatgcacatitggagaatgag	3961
Db	3901	cttcgcacatgagatggaatgaggtcgttgatgaataatagttgagatgcacatitggagaatgag	3960
QY	3962	agtaaaagaatatataagatgtgcacagatthactgcttataccctatttacaataatgtttgcc	4021
Db	3961	agtaaaagaatatataagatgtgcacagatthactgcttataccctatttacaataatgtttgcc	4020
QY	4022	acctggtctccttcaaaaagctcgacttcaactcccaaggagagtaatgctctgagacct	4081
Db	4021	acctggtctccttcaaaaagctcgacttcaactcccaaggagagtaatgctctgagacct	4080
QY	4082	caggttgataataatccaanaagggtgcgtgcgaatgagatcttccaagaagaacatataaatgaca	4141
Db	4081	caggttgataataatccaanaagggtgcgtgcgaatgagatcttccaagaagaacatataaatgaca	4140
QY	4142	ggagtaacctactcagggaatgaataatctctgcttaccagaatgtatgtaagagagttccctc	4201



```

Db      4141  ggaagtaactacgagggagtaaaatctctgtaccagcatgtatgtgaaggagttccctc 4200
Oy      4202  atccccagcagtcgaagatgagccatcagtcgagctctctttttcagaatgcaagtaag 4261
Db      4201  atccccagcagtcgaagatgagccatcagtcgagctctctttttcagaatgcaagtaag 4260
Oy      4262  gttttcagggaaatcaagaactcttcacacctgtgtgtgaactctcttaaccaccgtta 4321
Db      4261  gttttcagggaaatcaagaactcttcacacctgtgtgtgaactctcttaaccaccgtta 4320
Oy      4322  ctgactcgtctccttcgaattcaccgccagagttggtgtgacacaatgtgcctgagatg 4381
Db      4321  ctgactcgtctccttcgaattcaccgccagagttggtgtgacacaatgtgcctgagatg 4380
Oy      4382  gaggtctgggctcgcgaagcacaagacctctactga 4417
Db      4381  gaggtctgggctcgcgaagcacaagacctctactga 4416

RESULT 2
AA73164
ID  AA73164 standard; CDNA: 9354 BP.
XX
AC  AA73164;
XX
DT  08-APR-1998 (first entry)
XX
DE  cDNA encoding human B-domain deleted factor VIII.
XX
KW  Post-translational regulatory element; PRE; enhancer II; intronless gene;
KW  surface antigen gene; cytoplasmic accumulation; targeted delivery;
KW  near consensus splice sequence; blood coagulation factor; factor VIII;
KW  factor IX; ss.
XX
OS  Homo sapiens.
XX
XX
FH  Key
FH  CDS
FT  Location/Qualifiers
FT  2965..7380
FT  /*tag= a
FT  5165..5174
FT  /*tag= b
FT  /note= "5' near consensus site"
FT  5695..5703
FT  /*tag= c
FT  /note= "5' near consensus site"
FT  6320..6328
FT  /*tag= d
FT  /note= "5' near consensus site"
FT  6595..5603
FT  /*tag= e
FT  /note= "5' near consensus site"
FT  7045..7053
FT  /*tag= f
FT  /note= "5' near consensus site"
FT  7143..7152
FT  /*tag= g
FT  /note= "5' near consensus site"
FT  3296..3312
FT  /*tag= h
FT  /note= "3' near consensus site"
FT  4798..4817
FT  /*tag= i
FT  /note= "3' near consensus site"
FT  5023..5045
FT  /*tag= j
FT  /note= "3' near consensus site"
FT  5333..5355
FT  /*tag= k
FT  /note= "3' near consensus site"
FT  5520..5538
FT  /*tag= l
FT  /note= "3' near consensus site"

```

```

FT  misc_feature 5604..5632
FT  /*tag= m
FT  /note= "3' near consensus site"
FT  misc_feature 5717..5745
FT  /*tag= n
FT  /note= "3' near consensus site"
FT  misc_feature 6239..6258
FT  /*tag= o
FT  /note= "3' near consensus site"
FT  misc_feature 6658..6682
FT  /*tag= p
FT  /note= "3' near consensus site"
FT  misc_feature 7159..7176
FT  /*tag= q
FT  /note= "3' near consensus site"
FT  misc_feature 7196..7209
FT  /*tag= r
FT  /note= "3' near consensus site"
FT  misc_feature 7289..7315
FT  /*tag= s
FT  /note= "3' near consensus site"
FT  misc_feature 7411..7429
FT  /*tag= t
FT  /note= "3' near consensus site"
FT  misc_feature 7611..8197
FT  /*tag= u
FT  /note= "PRE sequence"

W09733994-A1.
18-SEP-1997.
10-MAR-1997; 97WO-US03561.
11-MAR-1996; 96US-0683839.
(IMMU-) IMMUNE RESPONSE CORP.
Bidingmaier S, ILL CR.
WPI, 1997-470874/43.
P-PSDB; AAW23414.

Vector for increased expression of intronless genes - comprises
intronless gene with at least one near consensus splice sequence, a
promoter and at least one viral cis-acting post-transcriptional
regulatory element

Example 1; Pages 21-31; 59pp; English.

The present sequence represents human B-domain deleted factor VIII
cDNA, and a post-translational regulatory element (PRE) of the
Hepatitis B virus, which is present 3' of the stop codon for factor VIII.
PRE sequences have been shown to function in cis to increase the
steady-state levels of surface gene transcripts by facilitating
cytoplasmic accumulation of these transcripts. The present sequence
is part of a novel vector, comprising an intronless gene containing
1 or more near consensus splice sequences operably linked to a
promoter sequence so that the gene is transcribed in a cell.
Intronless gene transcripts which contain near consensus splice site
sequences are believed to get tied up in the nucleus of the cell where
splicing occurs, rather than being transported to the cytoplasm where
they can be translated into proteins. The PRE sequences are transcribed
along with the gene, causing export of the gene transcript from the
nucleus into the cytoplasm of the cell. The vector can be used
to increase the expression of an intronless gene containing at least one
near consensus splice sites, preferably cDNA encoding a blood coagulation
factor, particularly factor VII or IX. The complex allows the targeted
delivery of the vector to a specific cell, e.g. hepatocytes when the
ligand is an asialoglycoprotein which binds the asialoglycoprotein
receptor present on their surface.

Sequence 9354 BP; 2506 A; 2239 C; 2161 G; 2448 T; 0 other;

```



Query Match	99.8%;	Score 4407.4;	DB 18;	Length 9354;
Best Local Similarity	99.9%;	Pred. No. 0;		
Matches 4411: Conservative	0;	Mismatches	6;	Indels 0; Gaps 0;

QY	1	caacaaatagaagctccacacgcgtctctctctgagccttttgagattcgtcttaagtc	60
Dp	2364	caatggtcttaagactccacacgcgtctctctctgagccttttgagattcgtcttaagtc	3023
QY	61	caccagaagaatactacctggtgtagcagtggaactgctacatgagactatctgcaagaatgtagct	120
Dp	3024	caccagaagaatactacctggtgtagcagtggaactgctacatgagactatctgcaagaatgtagct	3083
QY	181	cacctctgctgctgtacaaaagaagctctgtttttaaatcaagctgtacaccttttcaaat	240
Dp	3144	cacctctgctgctgtacaaaagaagctctgtttttaaatcaagctgtacaccttttcaaat	3203
QY	241	cgcataagaccagaagccacccctgtagtggtctgtctgaagctccacacacagcctgaggttta	300
Dp	3204	cgcataagaccagaagccacccctgtagtggtctgtctgaagctccacacacagcctgaggttta	3263
QY	301	tgatacagtggtcattacacacttaagaacaatggtctccacatccctgtcagttcttcagt	360
Dp	3264	tgatacagtggtcattacacacttaagaacaatggtctccacatccctgtcagttcttcagt	3323
QY	361	tgtgtgtatccctacttgnaaaagctctctgtaggaagcttaataitgataitgataagaccatcaag	420
Dp	3324	tgtgtgtatccctacttgnaaaagctctctgtaggaagcttaataitgataitgataagaccatcaag	3383
QY	421	ggagaaagaagaatcgtataaagctctccctctggtgtagaagccataatagtctgcaggtctct	480
Dp	3384	ggagaaagaagaatcgtataaagctctccctctggtgtagaagccataataatagtctgcaggtctct	3443
QY	481	gaaagaggaataggtcccaatggtcctctggaacccactgtgcttaacctatatactttctca	540
Dp	3444	gaaagaggaataggtcccaatggtcctctggaacccactgtgcttaacctatatactttctca	3503
QY	541	tgtggaacctgtataaagaacttgaattcagaagcctctctggaagccatacagatagtataga	600
Dp	3504	tgtggaacctgtataaagaacttgaattcagaagcctctctggaagccatacagatagtataga	3563
QY	601	agggaagtctgcccagaagaagaacacagaacctctgcacaattatatactacttttgcgt	660
Dp	3564	agggaagtctgcccagaagaagaacacagaacctctgcacaattatatactacttttgcgt	3623
QY	661	atttgtatgaaggnaaaagtgtgacctccagaaacaaagaacctctctgtatgcagagtaaga	720
Dp	3624	atttgtatgaaggnaaaagtgtgacctccagaaacaaagaacctctctgtatgcagagtaaga	3683
QY	721	tgtctgcatctgcgtcgagcctctgagccttaaaatgtagacacagtaaaatgtttaatgaacagctc	780
Dp	3684	tgtctgcatctgcgtcgagcctctgagccttaaaatgtagacacagtaaaatgtttaatgaacagctc	3743
QY	781	tctgccaagtgctgtagttgtagatgcacacagaagaatcagttcatattgtagcagtgtatgtgaatg	840
Dp	3744	tctgccaagtgctgtagttgtagatgcacacagaagaatcagttcatattgtagcagtgtatgtgaatg	3803
QY	841	caccacacctccgaagtgacatctcaatattcctctgnaaagttaacacatttctgtgnaagaacca	900
Dp	3804	caccacacctccgaagtgacatctcaatattcctctgnaaagttaacacatttctgtgnaagaacca	3863
QY	901	tcgcacagtgctctcttgnaaatctgcgcacataaacttctcttaactgctgcacaaacacctgtat	960
Dp	3864	tcgcacagtgctctcttgnaaatctgcgcacataaacttctcttaactgctgcacaaacacctgtat	3923
QY	961	ggacacttggaacagttctctactgtttttgtgcatactcctctccacacacacatgtagcatgga	1020
Dp	3924	ggacacttggaacagttctctactgtttttgtgcatactcctctccacacacacatgtagcatgga	3983

QY	1021	agcttatgtccaagtaagacagctgtccagaggaaccccaactacgaatgaaataatgta	1080
Db	3984	agcttatgtccaagtaagacagctgtccagaggaaccccaactacgaatgaaataatgta	4043
QY	1081	agaagcggagaagccttgaatgatacttaccgattctgaattgataatgtagtggcagaattgta	1140
Db	4044	agaagcggagaagccttgaatgatacttaccgattctgaattgataatgtagtggcagaattgta	4103
QY	1141	tgaatgaacatctccctcccttataccaatccgctcagttgcaccaagaagcatcctaaac	1200
Db	4104	tgaatgaacatctccctcccttataccaatccgctcagttgcaccaagaagcatcctaaac	4163
QY	1201	tgggttaacttaacttggctgtcgtgaaggaaggagactgggactatgctcccttagccctgc	1260
Db	4164	tgggttaacttaacttggctgtcgtgaaggaaggagactgggactatgctcccttagccctgc	4223
QY	1261	cccgcatacagaagttalaagaatcaatatttgacaatgycctccagcggattgtgag	1320
Db	4224	cccgcatacagaagttalaagaatcaatatttgacaatgycctccagcggattgtgag	4283
QY	1321	gaagtaacaaaaaagtcgcgatttatgtcatcaacagatgaaacctttaagactgtgaagc	1380
Db	4284	gaagtaacaaaaaagtcgcgatttatgtcatcaacagatgaaacctttaagactgtgaagc	4343
QY	1381	tattcagaatcgaatcagaatcccttggagccttacttattatgggaagtgttgagacaact	1440
Db	4344	tattcagaatcgaatcagaatcccttggagccttacttattatgggaagtgttgagacaact	4403
QY	1441	gttgattatatttaagaatccaagcagaagacatataactaactacccctcaggaattac	1500
Db	4404	gttgattatatttaagaatccaagcagaagacatataactaactacccctcaggaattac	4463
QY	1501	tgaatccgctcccttgatttccaaggaagtctccaagaagtataaacctttgaaggattt	1560
Db	4464	tgaatccgctcccttgatttccaaggaagtctccaagaagtataaacctttgaaggattt	4523
QY	1561	tccaattctgcagaagagaataattccaatataaatgycagtgactgtagaagatgggcc	1620
Db	4524	tccaattctgcagaagagaataattccaatataaatgycagtgactgtagaagatgggcc	4583
QY	1621	aactaatacagaatccctggctgtgcacccgcctataactctggttctgtaatatgggaag	1680
Db	4584	aactaatacagaatccctggctgtgcacccgcctataactctggttctgtaatatgggaag	4643
QY	1681	agatcagcttccaaggaacccatctgycctcctcctcactccgcacaaaagaatctgtagatca	1740
Db	4644	agatcagcttccaaggaacccatctgycctcctcctcactccgcacaaaagaatctgtagatca	4703
QY	1741	aagaggaacccaagataatgtccaagagaagaatgtcatccctgtttctgtattatgtatga	1800
Db	4704	aagaggaacccaagataatgtccaagagaagaatgtcatccctgtttctgtattatgtatga	4763
QY	1801	gaaccgaagctgtgatccctcacagaagaataatacaagctttctcccaatccaagcttggagt	1860
Db	4764	gaaccgaagctgtgatccctcacagaagaataatacaagctttctcccaatccaagcttggagt	4823
QY	1861	gcagctttgagatccaaggttccaagcctccaacatcatgtcacagatcalcaatgtgctatgt	1920
Db	4824	gcagctttgagatccaaggttccaagcctccaacatcatgtcacagatcalcaatgtgctatgt	4883
QY	1921	ttttgatagtttgcagttgttcaagtttgtttgacatgaggttgacataactggtatcttaag	1980
Db	4884	ttttgatagtttgcagttgttcaagtttgtttgacatgaggttgacataactggtatcttaag	4943
QY	1981	catltagaacacagatcagctccctctcgtctctctcctcctcgtatataccttcaaacaaa	2040
Db	4944	catltagaacacagatcagctccctctcgtctctctcctcctcgtatataccttcaaacaaa	5003
QY	2041	aatggtctatgaagaacacactcaaccctatactccaattctcagggaaactgtctcatgctc	2100
Db	5004	aatggtctatgaagaacacactcaaccctatactccaattctcagggaaactgtctcatgctc	5063



QY 2101 gatggaaacccaggtctatgatcttgggtgtccacaactcagaacttcggaacagag 2160  
|||||  
Db 5064 gatggaaacccaggtctatgatcttgggtgtccacaactcagaacttcggaacagag 5123  
QY 2161 catgaccgcttactgaagttcttaagtgtgtgaagaagaactggtgtgttttaagaga 2220  
|||||  
Db 5124 catgaccgcttactgaagttcttaagtgtgtgaagaagaactggtgtgttttaagaga 5183  
QY 2221 cagttcgaagatacttcagacatacttctgaatgaataaacagatcgcatctgaacagag 2280  
|||||  
Db 5184 cagttcgaagatacttcagacatacttctgaatgaataaacagatcgcatctgaacagag 5243  
QY 2281 ctctcccaagaatlcagaacacccctagacactagcagaagaactttaatgtccaccac 2340  
ctctcccaagaatlcagaacacccctagacactagcagaagaactttaatgtccaccac 5303  
Db 5244 ctctcccaagaatlcagaacacccctagacactagcagaagaactttaatgtccaccac 5303  
QY 2341 agtcttgaacagccataacagcggaataactcgttactactcttcagtcagatcaagaga 2400  
|||||  
Db 5304 agtcttgaacagccataacagcggaataactcgttactactcttcagtcagatcaagaga 5363  
QY 2401 aattgaatcagatataccataccatcagttgaatgaagaagaatttcgacattatga 2460  
|||||  
Db 5364 aattgaatcagatataccataccatcagttgaatgaagaagaatttcgacattatga 5423  
QY 2461 tgaagatgaaatlcagaagcccccgacgcttcaaaagaacacacacatattatgtc 2520  
|||||  
Db 5424 tgaagatgaaatlcagaagcccccgacgcttcaaaagaacacacacatattatgtc 5483  
QY 2521 tgaagatgaaaggtctggggttattgggattagctcccaactgttcttaagaacag 2580  
|||||  
Db 5484 tgaagatgaaaggtctggggttattgggattagctcccaactgttcttaagaacag 5543  
QY 2581 ggtcagaagtggcaggtgtccctcagttcaagaagaagtgttcttcagaagaatttaccagtg 2640  
|||||  
Db 5544 ggtcagaagtggcaggtgtccctcagttcaagaagaagtgttcttcagaagaatttaccagtg 5603  
QY 2641 ctctcttactcagcccttatacgttggagaactaaatgaacatttgggactctgggccc 2700  
ctctcttactcagcccttatacgttggagaactaaatgaacatttgggactctgggccc 5663  
Db 5604 ctctcttactcagcccttatacgttggagaactaaatgaacatttgggactctgggccc 5663  
QY 2701 ataatagaagcagaagtttgaagaataatacattgaacttcaagaataccagccctcgcg 2760  
|||||  
Db 5664 ataatagaagcagaagtttgaagaataatacattgaacttcaagaataccagccctcgcg 5723  
QY 2761 tccctatctctctcattctcactatcttctatagaagaatccagaagcagaagcaga 2820  
|||||  
Db 5724 tccctatctctctcattctcactatcttctatagaagaatccagaagcagaagcaga 5783  
QY 2821 acctagaanaaacttgcacagcctaatalgaaaccaaacttacttctggaaagtgcaca 2880  
acctagaanaaacttgcacagcctaatalgaaaccaaacttacttctggaaagtgcaca 5843  
Db 5784 acctagaanaaacttgcacagcctaatalgaaaccaaacttacttctggaaagtgcaca 5843  
QY 2881 tcatatggcccccttaagaatggattgacttgcacaaagcctggggttttctcagatgt 2940  
|||||  
Db 5844 tcatatggcccccttaagaatggattgacttgcacaaagcctggggttttctcagatgt 5903  
QY 2941 tgaacctgaaaaaagtgtgcctcagcgcttattgagaccccttcggtcgcacactaa 3000  
|||||  
Db 5904 tgaacctgaaaaaagtgtgcctcagcgcttattgagaccccttcggtcgcacactaa 5963  
QY 3001 cacactgaaccctgtcactcagtggaagaatgacagttacagaaatttgcctgttttccac 3060  
cacactgaaccctgtcactcagtggaagaatgacagttacagaaatttgcctgttttccac 6023  
Db 5964 cacactgaaccctgtcactcagtggaagaatgacagttacagaaatttgcctgttttccac 6023  
QY 3061 catctttgattgagaccaaagctgttacttcaactgaaataatgtgaaaaaacttgagggc 3120  
catctttgattgagaccaaagctgttacttcaactgaaataatgtgaaaaaacttgagggc 6083  
Db 6024 catctttgattgagaccaaagctgttacttcaactgaaataatgtgaaaaaacttgagggc 6083  
QY 3121 tccctcgaatatccagatgagagatcccaactttaaagaagatatcccttccatgcagat 3180  
|||||  
Db 6084 tccctcgaatatccagatgagagatcccaactttaaagaagatatcccttccatgcagat 6143  
QY 3181 caatgctacataatgatatcactactgtgttgaatggtcgaagatcaaaagattcg 3240  
|||||

Db 6144 caatgctacataatgatatcactactcgtgttgaatggtcgaagatcaaaagattcg 6203  
QY 3241 atggtatctgtccagatcgggcaggaatgaataatccatcttacttcaagtgcaga 3300  
atggtatctgtccagatcgggcaggaatgaataatccatcttacttcaagtgcaga 6263  
Db 6204 atggtatctgtccagatcgggcaggaatgaataatccatcttacttcaagtgcaga 6263  
QY 3301 tgtgtcactgttacgaaaaaaggaggtataaaatgagcactgtacatccactccacag 3360  
tgtgtcactgttacgaaaaaaggaggtataaaatgagcactgtacatccactccacag 6323  
Db 6264 tgtgtcactgttacgaaaaaaggaggtataaaatgagcactgtacatccactccacag 6323  
QY 3361 tgttttgaagacagtgtgaatgttaccatccaaagctggaatttggcgggtgtgaatgct 3420  
tgttttgaagacagtgtgaatgttaccatccaaagctggaatttggcgggtgtgaatgct 6383  
Db 6324 tgttttgaagacagtgtgaatgttaccatccaaagctggaatttggcgggtgtgaatgct 6383  
QY 3421 tatgtggagacatcaactgtctggatgagacacatttcttgggtgtacagcaataatgt 3480  
tatgtggagacatcaactgtctggatgagacacatttcttgggtgtacagcaataatgt 6443  
Db 6384 tatgtggagacatcaactgtctggatgagacacatttcttgggtgtacagcaataatgt 6443  
QY 3481 tcaagatccctcgggaatgtgtctcggacacatcagagatttccagatttcaagcttcag 3540  
tcaagatccctcgggaatgtgtctcggacacatcagagatttccagatttcaagcttcag 6503  
Db 6444 tcaagatccctcgggaatgtgtctcggacacatcagagatttccagatttcaagcttcag 6503  
QY 3541 acaatatgacagtgggcccccaagcttgcagacattcaattatcccgatccaatcagtc 3600  
acaatatgacagtgggcccccaagcttgcagacattcaattatcccgatccaatcagtc 6563  
Db 6504 acaatatgacagtgggcccccaagcttgcagacattcaattatcccgatccaatcagtc 6563  
QY 3601 ctggagacccaagggagcccttctcttgatcaagaatgtgattgtgtgcacaaatgatct 3660  
ctggagacccaagggagcccttctcttgatcaagaatgtgattgtgtgcacaaatgatct 6623  
Db 6564 ctggagacccaagggagcccttctcttgatcaagaatgtgattgtgtgcacaaatgatct 6623  
QY 3661 tcaacgcatcaagaacccaggtgtcccgctcagaagttctccagcctctacatctcagtt 3720  
tcaacgcatcaagaacccaggtgtcccgctcagaagttctccagcctctacatctcagtt 6683  
Db 6624 tcaacgcatcaagaacccaggtgtcccgctcagaagttctccagcctctacatctcagtt 6683  
QY 3721 tacaatcattatgattcttcttgggaagaatgtgcagacttatcgaagaatttccactgt 3780  
tacaatcattatgattcttcttgggaagaatgtgcagacttatcgaagaatttccactgt 6743  
Db 6684 tacaatcattatgattcttcttgggaagaatgtgcagacttatcgaagaatttccactgt 6743  
QY 3781 aaacttaatgtctctcttcttggcaatgtgattcactcctgggataaacaacatatctttaa 3840  
aaacttaatgtctctcttcttggcaatgtgattcactcctgggataaacaacatatctttaa 3840  
Db 6744 aaacttaatgtctctcttcttggcaatgtgattcactcctgggataaacaacatatctttaa 6803  
QY 3841 cccctcaattatgtctcgatataccggttgcacccaactcaattatagcatctgcagcac 3900  
cccctcaattatgtctcgatataccggttgcacccaactcaattatagcatctgcagcac 6863  
Db 6804 cccctcaattatgtctcgatataccggttgcacccaactcaattatagcatctgcagcac 6863  
QY 3901 tcttcgcatggagttgtatgggtgtgtatataatgttgcagcatctggaatgaga 3960  
tcttcgcatggagttgtatgggtgtgtatataatgttgcagcatctggaatgaga 6923  
Db 6864 tcttcgcatggagttgtatgggtgtgtatataatgttgcagcatctggaatgaga 6923  
QY 3961 gagtaagaacatatcagaatgacagatcactgtcacttccacttcaaatatgtttgc 4020  
gagtaagaacatatcagaatgacagatcactgtcacttccacttcaaatatgtttgc 6983  
Db 6924 gagtaagaacatatcagaatgacagatcactgtcacttccacttcaaatatgtttgc 6983  
QY 4021 caactgtgtctccttcaaaagctgactcactcaccctcaagggaggaatgactcgtgagacc 4080  
caactgtgtgtccttcaaaagctgactcactcaccctcaagggaggaatgactcgtgagacc 7043  
Db 6984 caactgtgtgtccttcaaaagctgactcactcaccctcaagggaggaatgactcgtgagacc 7043  
QY 4081 tcaagttgaataatccaaaagggtgtgcagatgtgacttccagaagaacatgaaagtccac 4140  
tcaagttgaataatccaaaagggtgtgcagatgtgacttccagaagaacatgaaagtccac 7103  
Db 7044 tcaagttgaataatccaaaagggtgtgcagatgtgacttccagaagaacatgaaagtccac 7103  
QY 4141 aggaatcaactcagagagtaaatcctcgtcttaccagcagtgtatgtgaaggttccct 4200  
aggaatcaactcagagagtaaatcctcgtcttaccagcagtgtatgtgaaggttccct 7163  
Db 7104 aggaatcaactcagagagtaaatcctcgtcttaccagcagtgtatgtgaaggttccct 7163  
QY 4201 catctccagcatcagaatggtccatcagttgagatctcttcttccagaatggaagaataa 4260  
catctccagcatcagaatggtccatcagttgagatctcttcttccagaatggaagaataa 7223  
Db 7164 catctccagcatcagaatggtccatcagttgagatctcttcttccagaatggaagaataa 7223  
QY 4261 ggttttcaaggaataatcagaactccttcaacacgtgtgtgaaactcttgaacccaagtt 4320  
ggttttcaaggaataatcagaactccttcaacacgtgtgtgaaactcttgaacccaagtt 4320  
|||||



Db 7224 gttttcagggaatacaagactcttcaacacctgtgtgtaactctctagaaccacggt 7283  
QY 4321 actgactcgcctacttgaaattacccccagagttgtgacacagattgcccgagat 4380  
Db 7284 actgactcgcctacttgaaattacccccagagttgtgacacagattgcccgagat 7343  
QY 4381 ggaagttctggcgtcgacagacagagactctactga 4417  
Db 7344 ggaagttctggcgtcgacagacagagactctactga 7380

RESULT 3  
AA076016  
ID AA076016 standard; cDNA; 4629 BP.  
XX  
AC AA076016;  
XX  
DT 20-JUL-1995 (first entry)  
XX  
DE B-domain deleted Factor-VIII.  
XX  
KW Factor-VIII; blood-clotting; hemophilia A; gene therapy;  
KW adenovirus; vector; ss.  
XX  
OS Homo sapiens.  
XX  
PN M09429471-A.  
XX  
PD 22-DEC-1994.  
XX  
PF 13-APR-1994; 94MO-US04075.  
XX  
PR 10-JUN-1993; 93US-0074920.  
PR 25-MAR-1994; 94US-0218353.  
XX  
PA (GENE-) GENETIC THERAPY INC.  
XX  
PI Connelly S, Kaleko M, Smith T;  
XX  
DR WPI; 1995-036495/05.  
DR P-PSDB; AAR67709.  
XX  
PT New adenoviral vectors for treatment of haemophilia - contg. a  
PT DNA sequence encoding a clotting factor, partic. Factor VIII or  
PT Factor IX  
XX  
PS Disclosure; Fig. 17A-17C; 116pp; English.  
XX  
CC Human Factor-VIII cDNA, from which the B domain had been deleted, was  
CC used to construct recombinant adenovirus vectors that produced  
CC therapeutic levels of the clotting factor when administered to an  
CC animal host, potentially providing hemophilia A gene therapy.  
XX  
SQ Sequence 4629 BP; 1318 A; 1051 C; 1018 G; 1242 T; 0 other;

Query Match 98.6%; Score 4355.2; DB 16; Length 4629;  
Best Local Similarity 99.1%; Pred. No. 0;  
Matches 4378; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 2 atcgaataagaagctccaccctgtcttctgtgaccttggcattgtcttagagcc 61  
Db 1 atcgaataagaagctccaccctgtcttctgtgaccttggcattgtcttagagcc 60  
QY 62 accagaagatactactcgtggtgtgcaatgtcatggaactatataatgcaagtgatctc 121  
Db 61 accagaagatactactcgtggtgtgcaatgtcatggaactatataatgcaagtgatctc 120  
QY 122 ggtgagactgctgtgagcgaagaattccctccctagagtgcccaaatctttcaatcaac 181  
Db 121 ggtgagactgctgtgagcgaagaattccctccctagagtgcccaaatctttcaatcaac 180  
QY 182 acccagctggtgtaaaaaaagactctgttctgtaaatcaacggttcaaccttcaacatc 241

Db 181 acccagctggtgtaaaaaaagactctgttctgtaaatcaacggttcaaccttcaacatc 240  
QY 242 gctaagccaagccaccctggtatgtctgtgaagtcctacacacagagtgagttat 301  
Db 241 gctaagccaagccaccctggtatgtctgtgaagtcctacacacagagtgagttat 300  
QY 302 gatacagttggtcattacacttaagaacatgcttcccatctgtccaattctcaagcttc 361  
Db 301 gatacagttggtcattacacttaagaacatgcttcccatctgtccaattctcaagcttc 360  
QY 362 ggtgtaccctactggaagactctcgaagagctgaatatgatagtatcgaaccagcaag 421  
Db 361 ggtgtaccctactggaagactctcgaagagctgaatatgatagtatcgaaccagcaag 420  
QY 422 gagaagaagatgataaagctctccctgtgtgaagacatacatatgtctgcaagctctg 481  
Db 421 gagaagaagatgataaagctctccctgtgtgaagacatacatatgtctgcaagctctg 480  
QY 482 aaagaagatgtgccaatgacctgaccacctgtgaccttaccctacatcatcttccat 541  
Db 481 aaagaagatgtgccaatgacctgaccacctgtgaccttaccctacatcatcttccat 540  
QY 542 gtgagactgtgtaaaagacttgaattcagacctcatctggaagccctagatgtagaagaa 601  
Db 541 gtgagactgtgtaaaagacttgaattcagacctcatctggaagccctacatactagaagaa 600  
QY 602 ggaagctgtgccaaggaagaacacagacctgtgacaatttatactactttgtctgta 661  
Db 601 ggaagctgtgccaaggaagaacacagacctgtgacaatttatactactttgtctgta 660  
QY 662 ttgtatgaagggaagaattgtgcactcagaacaaagaactccttgatgcagataagagat 721  
Db 661 ttgtatgaagggaagaattgtgcactcagaacaaagaactccttgatgcagataagagat 720  
QY 722 gctgcactgtctcggtcgccctgaataatgcacacagttcaattgttataaagctct 781  
Db 721 gctgcactgtctcggtcgccctgaataatgcacacagttcaattgttataaagctct 780  
QY 782 ctgccaagctcgtatgtgatgccaaggaatacagttcatatgtgcatgtgatatgagc 841  
Db 781 ctgccaagctcgtatgtgatgccaaggaatacagttcatatgtgcatgtgatatgagc 840  
QY 842 accactcctgaagtgaccttaatatctctcogaaggtcacacattcttgyagaagacat 901  
Db 841 accactcctgaagtgaccttaatatctctcogaaggtcacacattcttgyagaagacat 900  
QY 902 cgcagagctcctctggaatactcgcacaataacttcttactctgtcaaacacttgatg 961  
Db 901 cgcagagctcctctggaatactcgcacaataacttcttactctgtcaaacacttgatg 960  
QY 962 gaccttgaagattctactggtttgtcatatactctccaccacaatgatatgcatgaa 1021  
Db 961 gaccttgaagattctactggtttgtcatatactctctccaccacaatgatatgcatgaa 1020  
QY 1022 gcttatgtcaagtgagcagctgttccagaggaaccccaactagaatgaaataatgaa 1081  
Db 1021 gcttatgtcaagtgagcagctgttccagaggaaccccaactagaatgaaataatgaa 1080  
QY 1082 ggaagcgaagaactatgatatgtatcattacagattctgaatgtgagtgcaagttgat 1141  
Db 1081 ggaagcgaagaactatgatatgtatcattacagattctgaatgtgagtgcaagttgat 1140  
QY 1142 gatgacaactctccctcttatccaattcgtcagttgccaagaagacactcaaaact 1201  
Db 1141 gatgacaactctccctcttatccaattcgtcagttgccaagaagacactcaaaact 1200  
QY 1202 tgggtacattacattgtctgtgaagagagagctggaacatgctcccttagtctcgc 1261  
Db 1201 tgggtacattacattgtctgtgaagagagagctggaacatgctcccttagtctcgc 1260  
QY 1262 ccgagtgaagaagtataaagtcacataattgaacaattgcccctagcgagatgtgag 1321



Dh 1261 ccggaatgacagaagtltaataaagltcaataatttgaaacaaatgccccctcagcggatggttagg 1320  
Qy 1322 aagfacaaaaaagtlccggaattatagcatcacacagatgaaaccttaagaactgtaagct 1381  
Dh 1321 aagtaaaaaaaagtlccggaattatagcatcacacagatgaaaccttaagaactgtaagct 1380  
Qy 1382 atcagcatgaatcaagaactcttggaaccttaacttaatggggaagtctggagacaactg 1441  
Dh 1381 atcagcatgaatcaagaactcttggaaccttaacttaatggggaagtctggagacaactg 1440  
Qy 1442 ttgatttatattaagaatccaagcaagcagacacataaactaccctcagcaactcact 1501  
Dh 1441 ttgatttatattaagaatccaagcaagcagacacataaactaccctcagcaactcact 1500  
Qy 1502 gatctccgtcccttgatctcaagagatctaaccaaaagtgtlaaaacatttgaagaattt 1561  
Dh 1501 gatgtccgtcccttgatctcaagagatctaaccaaaagtgtlaaaacatttgaagaattt 1560  
Qy 1562 ccaattctgcagagagaataatccaataaavgacagatgactgtagaagaatgggcca 1621  
Dh 1561 ccaattctgcagagagaataatccaataaavgacagatgactgtagaagaatgggcca 1620  
Qy 1622 actaaatcagaatccctgggtgcctgagcccgctattacttaagtcttgaataatgagaga 1681  
Dh 1621 actaaatcagaatccctgggtgcctgagcccgctattacttaagtcttgaataatgagaga 1680  
Qy 1682 gatctaagcttcaggaactatctgagcccttcctcatctgtacaagaactctgtaagatca 1741  
Dh 1681 gatctaagcttcaggaactatctgagcccttcctcatctgtacaagaactctgtaagatca 1740  
Qy 1742 agaggaacacagataatctcagacaagaagaatgltcaactctgtttctgtatctgatatg 1801  
Dh 1741 agaggaacacagataatctcagacaagaagaatgltcaactctgtttctgtatctgatatg 1800  
Qy 1802 aacccaagctgtgtactcctcacagagaataataacaagctttctcccaatccagctgagtg 1861  
Dh 1801 aacccaagctgtgtactcctcacagagaataataacaagctttctcccaatccagctgagtg 1860  
Qy 1862 cagcttagagatccagagatcccaagcctccaacatcatgacagatcaatggtcatgtt 1921  
Dh 1861 cagcttagagatccagagatcccaagcctccaacatcatgacagatcaatggtcatgtt 1920  
Qy 1922 ttgtatagttgtcagatgtgtcagttgtttgtgcatgaggtgcatctgtgtacatctcaag 1981  
Dh 1921 ttgtatagttgtcagatgtgtcagttgtttcgtacatccactgtgcatctgtacatctcaag 1980  
Qy 1982 atgtgagacacagactgactccttctgtcttctctctgtgataactctcaacacaaa 2041  
Dh 1981 atgtgagacacagactgactccttctgtcttctctctgtgataactctcaacacaaa 2040  
Qy 2042 atgttctcatgaaagacacactacccatctccatctcagagagaacatggtctcatctg 2101  
Dh 2041 atgttctcatgaaagacacactacccatctccatctcagagagaacatggtctcatctg 2100  
Qy 2102 atggaaaaaccaggtctcatgtgattctgtgggtgtccacaactcagactctcgaaacagagc 2161  
Dh 2101 atggaaaaaccaggtctcatgtgattctgtgggtgtccacaactcagactctcgaaacagagc 2160  
Qy 2162 atgagcccttaactgaaaggtctctagttgtgacaagaacacagctggtatattacagagac 2221  
Dh 2161 atgagcccttaactgaaaggtctctagttgtgacaagaacacagctggtatattacagagac 2220  
Qy 2222 agtataagaatattctcagcatctgtctgtgataaaaaaaatgcatctgaaaccaagagc 2281  
Dh 2221 agtataagaatattctcagcatctgtctgtgataaaaaaaatgcatctgaaaccaagagc 2280  
Qy 2282 ttctctcagaatctcaagacacctagcaactagggcaaaaatgcaatcttaatgcaacccacca 2341  
Dh 2281 ttctctcagaatctcaagacacctagcaactagggcaaaaatgcaatcttaatgcaacccacca 2340  
Qy 2342 gtctctgaaacgcatcaacggaataaactcgtactactcttcagtctcaatgaagaa 2401  
Dh 2341 gtctctgaaacgcatcaacggaataaactcgtactactcttcagtctcaatgaagaa 2400

Qy 2402 atgacatgatgatcacatcatcagttgaaatgaaagaagaattttgacatttatgat 2461  
Dh 2401 atgacatgatgatcacatcatcagttgaaatgaaagaagaattttgacatttatgat 2460  
Qy 2462 gaggaatgaataatcagagcccccagacttccaagaagaacagacactatttatgtct 2521  
Dh 2461 gaggaatgaataatcagagcccccagacttccaagaagaacagacactatttatgtct 2520  
Qy 2522 gcaagtggagagctcttggaattatggaatggaatctccacatggtctcaagaagaag 2581  
Dh 2521 gcaagtggagagctcttggaattatggaatggaatctccacatggtctcaagaagaag 2580  
Qy 2582 gctcagagctgagaggtctccctcagttccaagaagaatgttttccaggaatttactcagc 2641  
Dh 2581 gctcagagctgagaggtctccctcagttccaagaagaatgttttccaggaatttactcagc 2640  
Qy 2642 tcccttactcagcccttaacccgtgagagaactaaatgacaatttgggaactctctgggcca 2701  
Dh 2641 tcccttactcagcccttaacccgtgagagaactaaatgacaatttgggaactctctgggcca 2700  
Qy 2702 tatataagagcagaagttgaaagataataatcatggttaacttcagaaatcagagctctcgt 2761  
Dh 2701 tatataagagcagaagttgaaagataataatcatggttaacttcagaaatcagagctctcgt 2760  
Qy 2762 cccatctcccttaattctagaccttaattcttaatgagaagaatcagagcgcaagagcagaa 2821  
Dh 2761 cccatctcccttaattctagaccttaattcttaatgagaagaatcagagcgcaagagcagaa 2820  
Qy 2822 cctagaaaaaactctgtcagaagcctaaatgaaaccaaactactctttgaaagtgtcaacat 2881  
Dh 2821 cctagaaaaaactctgtcagaagcctaaatgaaaccaaactactctttgaaagtgtcaacat 2880  
Qy 2882 catatggaccccaactaaagaatgagtttgaactgacagcagctgggttttctctcagatgt 2941  
Dh 2881 catatggaccccaactaaagaatgagtttgaactgacagcagctgggttttctctcagatgt 2940  
Qy 2942 gaccctgaaagaatgltgtcactcagcctgtatctggaacccctctcgtgtccacactaac 3001  
Dh 2941 gaccctgaaagaatgltgtcactcagcctgtatctggaacccctctcgtgtccacactaac 3000  
Qy 3002 acaactgaacccctgtcactatggtgagacaagtgtacagtlacaggaattgtctcgttttccac 3061  
Dh 3001 acaactgaacccctgtcactatggtgagacaagtgtacagtlacaggaattgtctcgttttccac 3060  
Qy 3062 atcttctgaatggaccaaagaactggtgtacttcaactgaataatggaagaagaactgtcagagct 3121  
Dh 3061 atcttctgaatggaccaaagaactggtgtacttcaactgaataatggaagaagaactgtcagagct 3120  
Qy 3122 cccctgcaatctcagatggaagaatcccaactttaaagaagaattatcgcttccatgtcaatc 3181  
Dh 3121 cccctgcaatctcagatggaagaatcccaactttaaagaagaattatcgcttccatgtcaatc 3180  
Qy 3182 aatggtctacataatggaatacactacctgtgttagtaatggtctcagatccaagaattcga 3241  
Dh 3181 aatggtctacataatggaatacactacctgtgttagtaatggtctcagatccaagaattcga 3240  
Qy 3242 tggatctgcgtcgaagcatggtggcagcaatgaaacatccaattcatctcatctcagtgagcat 3301  
Dh 3241 tggatctgcgtcagatggtggcagcaatgaaacatccaattcatctcatctcagtgagcat 3300  
Qy 3302 gtgttctactgttgaagaaaaaagagagatataaaatggtcactgtataactctcatccaggt 3361  
Dh 3301 gtgttctactgttgaagaaaaaagagagatataaaatggtcactgtataactctcatccaggt 3360  
Qy 3362 gtttttgagacatgvgaaatgttaaccatccaagcttggaaattvggctgggtgtgaatgcctt 3421  
Dh 3361 gtttttgagacatgvgaaatgttaaccatccaagcttggaaattvggctgggtgtgaatgcctt 3420  
Qy 3422 attggtgagcatctacatgctgggtgagcagcacttttctcgtgtacagcaataaagtgt 3481  
Dh 3421 attggtgagcatctacatgctgggtgagcagcacttttctcgtgtacagcaataaagtgt 3480



```

QY 3482 cagactccctgggaatgctctggaacacatagaattctcagatcagctcaagga 3541
    |||||||
Db 3481 cagactccctgggaatgctctggaacacatagaattctcagatcagctcaagga 3540
QY 3542 caatatgacagctgggcccacaagctggcagactcattatccggatcaataatgcc 3601
    |||||||
Db 3541 caatatgacagctgggcccacaagctggcagactcattatccggatcaataatgcc 3600
QY 3602 tggagaccacgaagagcccttcttggaatgaagtgagctgtctggcaccgaatgata 3661
    |||||||
Db 3601 tggagaccacgaagagcccttcttggaatgaagtgagctgtctggcaccgaatgata 3660
QY 3662 cagcgatcaagaagcagaggtgcccgtcagaagatctccagcctcacaatccagtt 3721
    |||||||
Db 3661 cagcgatcaagaagcagaggtgcccgtcagaagatctccagcctcacaatccagtt 3720
QY 3722 atcatcatgatagtcttgatgaggaagaagtggcagactatcgagggaattccactgga 3781
    |||||||
Db 3721 atcatcatgatagtcttgatgaggaagaagtggcagactatcgagggaattccactgga 3780
QY 3782 acccttaatgctctcttggaatgagatgattcatctgggataaacaacataattttaac 3841
    |||||||
Db 3781 acccttaatgctctcttggaatgagatgattcatctgggataaacaacataattttaac 3840
QY 3842 cctccaattatgctcgatatacatcgtttggaccacaactatataagatccgagcact 3901
    |||||||
Db 3841 cctccaattatgctcgatatacatcgtttggaccacaactatataagatccgagcact 3900
QY 3902 ctgcgacatgagatgagatggcgtgatttaaatagttgcaagatgcaattgggaatggag 3961
    |||||||
Db 3901 ctgcgacatgagatgagatggcgtgatttaaatagttgcaagatgcaattgggaatggag 3960
QY 3962 agtaagaacataatagatggcacaatatactgctcatcctccttaccataatgttggcc 4021
    |||||||
Db 3961 agtaagaacataatagatggcacaatatactgctcatcctccttaccataatgttggcc 4020
QY 4022 accggtctccttcaaaaagctcgactcaacctccaaggaaggaatagtccttgagacct 4081
    |||||||
Db 4021 accggtctccttcaaaaagctcgactcaacctccaaggaaggaatagtccttgagacct 4080
QY 4082 caggtgaataatccaaaagatggctgcaagtggactccagaagaacaatgaaatgcaaca 4141
    |||||||
Db 4081 caggtgaataatccaaaagatggctgcaagtggactccagaagaacaatgaaatgcaaca 4140
QY 4142 ggaataactactcaggaagtaaaatctcgttaccagcatgatagtggaaggaattcttc 4201
    |||||||
Db 4141 ggaataactactcaggaagtaaaatctcgttaccagcatgatagtggaaggaattcttc 4200
QY 4202 atcccgacgaatcagaatgagcatcagtgagactctcttttcagaatggcaagtaag 4261
    |||||||
Db 4201 atcccgacgaatcagaatgagcatcagtgagactctcttttcagaatggcaagtaag 4260
QY 4262 gttttcaagggaatcaagaatcctcacaacctggtgtaaatctctagaccacgltta 4321
    |||||||
Db 4261 gttttcaagggaatcaagaatcctcacaacctggtgtaaatctctagaccacgltta 4320
QY 4322 ctgactcgtaactctgtaattcaacccccaggttgggtgacccagatgtgccctgaagatg 4381
    |||||||
Db 4321 ctgactcgtaactctgtaattcaacccccaggttgggtgacccagatgtgccctgaagatg 4380
QY 4382 gaagttctggctgagagcacaagacactcactga 4417
    |||||||
Db 4381 gaagttctggctgagagcacaagacactcactga 4416

```

RESULT 4  
AAV23339  
ID AAV23339 standard: DNA: 4670 BP.

XX AAV23339:  
XX  
XX 17-AUG-1998 (first entry)  
XX

```

DE Human Factor-VIII gene lacking central B domain.
XX
KW Factor-VIII; blood clotting; human; synthetic gene; codon usage; ss.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WC09812207-A1.
XX
PD 26-MAR-1998.
XX
PE 18-SEP-1997; 97WO-US16639.
XX
PR 20-SEP-1996; 96US-0717294.
XX
PA (GEHO ) GEN HOSPITAL CORP.
XX
PI Haas J, Seed B;
XX
DR WPI; 1998-217200/19.
XX
XX New synthetic eukaryotic gene(s) - in which non-preferred or less
PT preferred codon(s) are replaced to provide high level expression in
PT mammalian cell(s)
XX
XX Example 3; Fig 12; 92pp; English.
XX
CC This gene codes for a human Factor-VIII protein that lacks the
CC central B domain (amino acids 760-1639) of the native protein. In
CC a novel, claimed synthetic gene (see AAV23388), non-preferred or
CC less preferred codons of the native gene are replaced by codons
CC favored by highly expressed human genes to provide high-level
CC expression in mammalian cells. The synthetic gene was assembled
CC from 29 pairs of oligonucleotides (see AAV23340-97) which served as
CC PCR templates. Synthetic genes of the invention (see also
CC AAV23289-91) are used for production of recombinant proteins in
CC mammalian cells at levels of at least 500% of those obtained using
CC the natural genes. They can also be used in gene therapy.
XX
SQ Sequence 4670 BP; 1325 A; 1072 C; 1036 G; 1237 T; 0 other;

Query Match 98.5%; Score 4351; DB 19; Length 4670;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 4354; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 59 gccaccagaatactactcctggtgcagtggaactgtcatggaatataatgcaagtat 118
    |||||||
Db 84 gccaccagaatactactcctggtgcagtggaactgtcatggaatataatgcaagtat 143
QY 119 ctgggtgagctgctgtggacgcaagattcctccttagagtgccaaatctttccattc 178
    |||||||
Db 144 ctgggtgagctgctgtggacgcaagattcctccttagagtgccaaatctttccattc 203
QY 179 aacccctcagtcggtacaacaaagactctgtttgaatatcaaggttcaccttttaac 238
    |||||||
Db 204 aacccctcagtcggtacaacaaagactctgtttgaatatcaaggttcaccttttaac 263
QY 239 atcgtaagccaagaagcaccctgagatggtctgtaagctctaccatccaggtcgaagt 298
    |||||||
Db 264 atcgtaagccaagaagcaccctgagatggtctgtaagctctaccatccaggtcgaagt 323
QY 299 tatgatacagtggtcatatacaacttaagaacatggtctccatcctgtcagttcattgct 358
    |||||||
Db 324 tatgatacagtggtcatatacaacttaagaacatggtctccatcctgtcagttcattgct 383
QY 359 gtgtgtatccctactggaagagctctcgaaggagctgaatatgatagtatcgaccagctcaa 418
    |||||||
Db 384 gtgtgtatccctactggaagagctctcgaaggagctgaatatgatagtatcgaccagctcaa 443
QY 419 agggagaagaagaatgaataagcttccctggttggaaagcacaataatgtctggcagctc 478
    |||||||
Db 444 agggagaagaagaatgaataagcttccctggttggaaagcacaataatgtctggcagctc 503

```



QY 479 ctgaagaagaa tgytccaatggcctctgacccactgcttactactacatcttct 538  
|||||  
Db 504 ctgaagaagaa tgytccaatggcctctgacccactgcttactactacatcttct 563  
QY 539 catgtgacactggttaaaaagacttgaattcagggcctcatcttgagccctactagtatgtaga 598  
564 catgtgacactggttaaaaagacttgaattcagggcctcatcttgagccctactagtatgtaga 623  
QY 599 gaaggagctctggccaagaaagacacagacacttgcacaaatttatactacttttctgt 658  
|||||  
Db 624 gaaggagctctggccaagaaagacacagacacttgcacaaatttatactacttttctgt 683  
QY 659 gtatttga tgaagggaagaaagtctgacatcagaacaaagaactctcttga tgcaggatag 718  
684 gtatttga tgaagggaagaaagtctgacatcagaacaaagaactctcttga tgcaggatag 743  
QY 719 gatgtctgacatctgtctgggctggtccctaataatgcacacagtcata tgytataacag 778  
744 gatgtctgacatctgtctgggctggtccctaataatgcacacagtcata tgytataacag 803  
QY 779 tctctgccaagttctgattggatctgcacaggaatacagttacttggcatgtatggaa t 838  
|||||  
Db 804 tctctgccaagttctgattggatctgcacaggaatacagttacttggcatgtatggaa t 863  
QY 839 ggcacacactctctgaagtgacatcaaatctcctcgaagg tcaacacattctcttggagac 898  
864 ggcacacactctctgaagtgacatcaaatctcctcgaagg tcaacacattctcttggagac 923  
QY 899 catctgcagagcgctctctggaatactcgcgaataacttccctactctgctcaaacctctg 958  
924 catctgcagagcgctctctggaataactcgcgaataacttccctactctgctcaaacctctg 983  
QY 959 atggaccttgacaggttctctactgttctgtcatctctctccacacaaatgacatg 1018  
|||||  
Db 984 atggaccttgacaggttctctactgttctgtcatctctctccacacaaatgacatg 1043  
QY 1019 gaagcttatgtcaaatgtagacagctgtccagaggaaccccaacta cgaatgaaaaataat 1078  
1044 gaagcttatgtcaaatgtagacagctgtccagaggaaccccaacta cgaatgaaaaataat 1103  
QY 1079 gaagaaacggaagacataatgataatcttactgtacttctgaaatgata tgytgcaggtt 1138  
1104 gaagaaacggaagacataatgataatcttactgtacttctgaaatgata tgytgcaggtt 1163  
QY 1139 gatgatcaaacactctcctctcttacttcaaatctgcctcagttgcaggaagaatccttaa 1198  
|||||  
Db 1164 gatgatcaaacactctcctctcttacttcaaatctgcctcagttgcaggaagaatccttaa 1223  
QY 1199 acttgggtacatatcatctgtctgtgaaggagagactggnactatgtctcttactctc 1258  
1224 acttgggtacatatcatctgtctgtgaaggagagactggnactatgtctcttactctc 1283  
QY 1259 gcccccga tgcacaaagttaaaaaagtcataattgaaacaatggccctcagcgagttggt 1318  
1284 gcccccga tgcacaaagttaaaaaagtcataattgaaacaatggccctcagcgagttggt 1343  
QY 1319 aggaagtaacaaaagtcgcgattta tggcatcacacagatgaactttaagactgttaa 1378  
1344 aggaagtaacaaaagtcgcgattta tggcatcacacagatgaactttaagactgttaa 1403  
QY 1379 gctattcagca tgaatcaggaatcttggaccttacttctatctggtggaagt tggagacca 1438  
1404 gctattcagca tgaatcaggaatcttggaccttacttctatctggtggaagt tggagacca 1463  
QY 1439 ctgtgtgattatatttaagaatccaagcaagcagacataatcaatctaccctcaggaatc 1498  
1464 ctgtgtgattatatttaagaatccaagcaagcagacataatcaatctaccctcaggaatc 1523  
QY 1499 actgatctcgtctcttcttattcaaggaatatacacaagaagtgtaaaacatttgaagat 1558  
1524 actgatctcgtctcttcttattcaaggaatatacacaagaagtgtaaaacatttgaagat 1583

QY 1559 ttccaatctctgcagaggaataattcaataataatgagacagtgtagaaatgag 1618  
|||||  
Db 1584 ttccaatctctgcagaggaataattcaataataatgagacagtgtagaaatgag 1643  
QY 1619 ccaactaaatcagatccctcggtgcttgacccgctattacttactagtttcgttaaatgag 1678  
1644 ccaactaaatcagatccctcggtgcttgacccgctattacttactagtttcgttaaatgag 1703  
QY 1679 agagatctagctctcagactcaatctggccctcccaactctcacaagtctgatat 1738  
1704 agagatctagctctcagactcaatctggccctcccaactctcacaagtctgatat 1763  
QY 1739 caaagaggaaacagataatg tlcagacaagaggaatgtcatccgtttctctgtattgtat 1798  
1764 caaagaggaaacagataatg tlcagacaagaggaatgtcatccgtttctctgtattgtat 1823  
QY 1799 gagaacccgaagctgtgtactctcagagataatacaacgcttctcccaatccagctgga 1858  
1824 gagaacccgaagctgtgtactctcagagataatacaacgcttctcccaatccagctgga 1883  
QY 1859 gtgcagcttgagatccaaggttccaagcctccaacatcatgacaacatcatgtgtat 1918  
1884 gtgcagcttgagatccaaggttccaagcctccaacatcatgacaacatcatgtgtat 1943  
QY 1919 gtitttga tgaatttgcagttgtcaagttgttgcatagtgagtgacatgtatactcta 1978  
1944 gtitttga tgaatttgcagttgtcaagttgttgcatagtgagtgacatgtatactcta 2003  
QY 1979 agcatltgagacaaagactgacttctctctgtcttcttctcttgatataacttcaaacac 2038  
2004 agcatltgagacaaagactgacttctctctgtcttcttctcttgatataacttcaaacac 2063  
QY 2039 aaaatggtctatgaaagacacactccaactatcccatctcaggagaacatgttctcatg 2098  
2064 aaaatggtctatgaaagacacactccaactatcccatctcaggagaacatgttctcatg 2123  
QY 2099 tcatgtgaaaaacccagctctcatgtgaattctgggtgtgcacaactcagaacttcggagaca 2158  
2124 tcatgtgaaaaacccagctctcatgtgaattctgggtgtgcacaactcagaacttcggagaca 2183  
QY 2159 ggcattgacgcgtcttacttgaagtttctgtgtgtgacaaagacatgtgtattatagag 2218  
2184 ggcattgacgcgtcttacttgaagtttctgtgtgtgacaaagacatgtgtattatagag 2243  
QY 2219 gacagttatgaagatatttcagacttctgtctggtlaaaaaaatgtcatgtgaacacaga 2278  
2244 gacagttatgaagatatttcagacttctgtctggtlaaaaaaatgtcatgtgaacacaga 2303  
QY 2279 agctctccccaatctcaagacacccctagcactagycacaaagcaaatttaatgccaacca 2338  
2304 agctctccccaatctcaagacacccctagcactagycacaaagcaaatttaatgccaacca 2363  
QY 2339 ccagctctgaaacggtccatcaacggaataactgttactactcttgcgtgaatcagaag 2398  
2364 ccagctctgaaacggtccatcaacggaataactgttactactcttgcgtgaatcagaag 2423  
QY 2399 gaaatctgactatgatacatatcaatcaatltgaaatgaaagagagaatttgcacattat 2458  
2424 gaaatctgactatgatacatatcaatcaatltgaaatgaaagagagaatttgcacattat 2483  
QY 2459 gatgagga tgaanaatcagagcccccgcagcttccaagaagaacagacataattat 2518  
2484 gatgagga tgaanaatcagagcccccgcagcttccaagaagaacagacataattat 2543  
QY 2519 gctgtcagtgagagagctcttggtattatgaggtgtgtgtgtcccaatggttttaagaac 2578  
2544 gctgtcagtgagagagctcttggtattatgaggtgtgtgtgtcccaatggttttaagaac 2603  
QY 2579 agggctcagagtgagaggtccctccaggtccaagaagaattgtttccaggaaatttaagat 2638  
2604 agggctcagagtgagaggtccctccaggtccaagaagaattgtttccaggaaatttaagat 2663  
QY 2639 ggtctcttactcagccctataccgttggagaactaaatgaaacttggagctctcggg 2698



```
Db 2664 ggcctcttactcagcccttactacgctgaggaactaataatgacaatttgggcctcttggg 2723
Qy 2699 ccatatataagagcagaagtgtgaagataataatcagttaacttcagaatcaggcctct 2758
Db 2724 ccatatataagagcagaagtgtgaagataataatcagttaacttcagaatcaggcctct 2783
Qy 2759 gcttcctattcctctctatttctagccttatttctataggaagatcagagcgcaagggca 2818
Db 2784 cgtccctattcctctctatttctagccttatttctataggaagatcagagcgcaagggca 2843
Qy 2819 gaacctagaataaaacttctgaagccttaataatgaaaccaaacttcttggaaagctgcaa 2878
Db 2844 gaacctagaataaaacttctgaagccttaataatgaaaccaaacttcttggaaagctgcaa 2903
Qy 2879 catcatatgacacccactaaagaatgtgaactgtcaagcctgagcttattctctgat 2938
Db 2904 catcatatgacacccactaaagaatgtgaactgtcaagcctgagcttattctctgat 2963
Qy 2939 gttgacctggaataaagatgtgcaactcagcctgaattggaacctctgagctgacacact 2998
Db 2964 gttgacctggaataaagatgtgcaactcagcctgaattggaacctctgagctgacacact 3023
Qy 2999 aacacactgaaacctgtctcatatgaggaacaaagtgaacagtaacaggaattgtctgttttc 3058
Db 3024 aacacactgaaacctgtctcatatgaggaacaaagtgaacagtaacaggaattgtctgttttc 3083
Qy 3059 accaactttatgtagacacaaagctgtgtaactcaatgaaatatgaaagaactgcagag 3118
Db 3084 accaactttatgtagacacaaagctgtgtaactcaatgaaatatgaaagaactgcagag 3143
Qy 3119 gctccctgcacatatccagatggaagatcccaacttttaagaagaatttcgtcttcacgca 3178
Db 3144 gctccctgcacatatccagatggaagatcccaacttttaagaagaatttcgtcttcacgca 3203
Qy 3179 atcaaatgtctacataatgagatatacctacactgagcttaagttaatgagctcaggaatc 3238
Db 3204 atcaaatgtctacataatgagatatacctacactgagcttaagttaatgagctcaggaatc 3263
Qy 3239 cgaatgtatctgtctcagcatgagcgagcaatgaaacatcattctatcttcaagtga 3298
Db 3264 cgaatgtatctgtctcagcatgagcgagcaatgaaacatcattctatcttcaagtga 3323
Qy 3299 catgtgtcactgtgacgaaataaagaagagatataaataatggaactgtaacatctaccca 3358
Db 3324 catgtgtcactgtgacgaaataaagaagagatataaataatggaactgtaacatctaccca 3383
Qy 3359 ggtgttttggagacagtgaataatgttaccatccaaagcttgaatttggcggtggaatgc 3418
Db 3384 ggtgttttggagacagtgaataatgttaccatccaaagcttgaatttggcggtggaatgc 3443
Qy 3419 ctatattggcagacatctacatgtctggaatgagcaacacttttctgtgtcacaagaataag 3478
Db 3444 ctatattggcagacatctacatgtctggaatgagcaacacttttctgtgtcacaagaataag 3503
Qy 3479 tgtcagaagctcccgagggaatgagcttctggaacacatlagagaattttagagcttca 3538
Db 3504 tgtcagaagctcccgagggaatgagcttctggaacacatlagagaattttagagcttca 3563
Qy 3539 ggaacaataatgagacagtggccccaagaagctggcaggaattctatattccggatcacaat 3598
Db 3564 ggaacaataatgagacagtggccccaagaagctggcaggaattctatattccggatcacaat 3623
Qy 3599 gccctggagacccaagagcccttcttctgtgatacaaggtgagatcgtgtggccaacatg 3658
Db 3624 gccctggagacccaagagcccttcttctgtgatacaaggtgagatcgtgtggccaacatg 3683
Qy 3659 attcacggcatcaagaccagcggtgcctgcagagaagttcccaagccctctacatctcag 3718
Db 3684 attcacggcatcaagaccagcggtgcctgcagagaagttcccaagccctctacatctcag 3743
Qy 3719 ttatcatcatgtatagtcttgaatgggaagaagtgcagacattacgagaagaattccact 3778
```

```
Db 3744 ttatcatcatgtatagtcttgaatgggaagaagtgcagacattacgagaagaattccact 3803
Qy 3779 ggaaccttaatgtgtcttctcttggcaatgtgatactcactcgtggataaacaacataatttc 3838
Db 3804 ggaaccttaatgtgtcttctcttggcaatgtgatactcactcgtggataaacaacataatttc 3863
Qy 3839 aaccttcaattattgtctgatacatcogtttggacccaactcaatttaacattcagcagc 3898
Db 3864 aaccttcaattattgtctgatacatcogtttggacccaactcaatttaacattcagcagc 3923
Qy 3899 actcttcgcatgagatgtgatacgtgagctgataatagattgcagacatgacatttggagatg 3958
Db 3924 actcttcgcatgagatgtgatacgtgagctgataatagattgcagacatgacatttggagatg 3983
Qy 3959 gagagtaagaacataatcagatgacaggaattacgtcttcaacttaccataatgatt 4018
Db 3984 gagagtaagaacataatcagatgacaggaattacgtcttcaacttaccataatgatt 4043
Qy 4019 gccacctggtctccttcaaaaagctcgactccaactccaagggaggaattagctctggaga 4078
Db 4044 gccacctggtctccttcaaaaagctcgactccaactccaagggaggaattagctctggaga 4103
Qy 4079 cctcaggtgataatcccaaaaaggtggtcgaagtgaacttccagagaacatgaaagtc 4138
Db 4104 cctcaggtgataatcccaaaaaggtggtcgaagtgaacttccagagaacatgaaagtc 4163
Qy 4139 acaggaagtaactactcagaggaatgaaatctctgcttaccagcatgtatgtgaagagttc 4198
Db 4164 acaggaagtaactactcagaggaatgaaatctctgcttaccagcatgtatgtgaagagttc 4223
Qy 4199 ctcaatccagcagcagtaagaattgacatcagtgagctctctttttcagaatggaagaatga 4258
Db 4224 ctcaatccagcagcagtaagaattgacatcagtgagctctctttttcagaatggaagaatga 4283
Qy 4259 aagggttttcaagggaataatcagaactccttcaacactgtgtgtaactctcagaaccacg 4318
Db 4284 aagggttttcaagggaataatcagaactccttcaacactgtgtgtaactctcagaaccacg 4343
Qy 4319 ttactgaactgcataccttctgaattcaaccacagagttgggtgacacagattgacctgaag 4378
Db 4344 ttactgaactgcataccttctgaattcaaccacagagttgggtgacacagattgacctgaag 4403
Qy 4379 atgagagttctgggctcgagagcgacagacctactga 4417
Db 4404 atgagagttctgggctcgagagcgacagacctactga 4442

RESULT 5
AAV19581
ID AAV19581 standard; cDNA; 4832 BP.
XX
AC AAV19581;
XX
DT 06-AUG-1998 (first entry)
XX
DE Human factor VIII beta-domain deleted SQN deletion cDNA sequence.
XX
KW Replication defective; recombinant retrovirus; RRV; therapeutic protein;
KW haemophilia; thrombosis; hypercoagulable disorder; liver disease; human;
KW hepatitis; thalassemia; phenylketonuria; Lesch-Nyhan syndrome; diabetes;
KW cystic fibrosis; Duchenne's Muscular Dystrophy; hypercholesterolemia;
KW hypopituitarism; adenine deaminase deficiency; HIV infection; anaemia;
KW Gucher's syndrome; high blood pressure; Alzheimer's disease; autoimmune;
KW inflammatory disease; factor VIII; ss.
XX
OS Homo sapiens.
XX
PN M09800541-AZ.
XX
PD 08-JAN-1998.
XX
PF 02-JUL-1997; 97MO-US11784.
XX
```



PR 04-JUN-1997; 97US-0869309.  
PR 03-JUL-1996; 96US-0645601.  
PR 13-AUG-1996; 96US-0696381.  
XX  
XX (CHIR ) CHIRON CORP.  
XX Allen JR, Barber JR, Boder M, Chang SMW, Chong K;  
PI De La Vega D, Depolon J, Greengard J, Hsu DC, Ibanez CE;  
PI Jolly DJ, Lee R, Mittelstaedt DM, Prussak CE, Respess JG;  
XX WPI: 1998-086966/08.  
DR P-PSDB; AAW46246.  
XX  
XX New replication defective recombinant retro-viruses - which can be  
PT administered to provide long term systemic expression of therapeutic  
PT protein in blood, useful in, e.g. treating hyper-coagulable  
PT disorders  
XX  
XX Example 28; Pages 210-213; 272pp; English.  
PS  
XX This CDNA encodes the beta-domain deleted SQN deletion protein of human  
CC factor VIII. This is used in the construction of recombinant retroviral  
CC vectors expressing human factor VIII. The invention provides the  
CC preparation of replication defective recombinant retrovirus (RVV)  
CC expressing a therapeutic protein. The RVV preparation is resistant to  
CC degradation by human complement and is capable of inducing long term  
CC systemic expression of the therapeutic protein when administered  
CC intravenously to a human. The long term systemic expression results in a  
CC measurable level of the therapeutic protein being produced in the blood  
CC of the human for a period of at least 30 days after the administration of  
CC the RVV vector preparation. RVV's can be used for in vivo delivery of  
CC therapeutic protein to treat, e.g. haemophilia A, haemophilia B,  
CC thrombosis, hypercoagulable disorders, liver diseases such as hepatitis,  
CC severe combined immunodeficiency (SCID), cystic fibrosis, Duchenne's  
CC muscular dystrophy, inherited emphysema, familial hypercholesterolemia,  
CC diabetes, hypopituitarism, adenine deaminase deficiency, alphas-  
CC antitrypsin deficiency, Gaucher's syndrome, anaemia, infections such as  
CC HIV infection, high blood pressure, Alzheimer's disease, autoimmune or  
CC inflammatory disease or graft versus host disease. RVV's are capable of  
CC surviving inactivation in human serum thereby allowing efficient gene  
CC transfer over prolonged periods of time.  
XX  
XX Sequence 4832 BP; 1381 A; 1089 C; 1055 G; 1307 T; 0 other;  
SQ  
Query Match 97.8%; Score 4321.4; DB 19; Length 4832;  
Best Local Similarity 99.0%; Pred. No. 0;  
Matches 4374; Conservative 0; Mismatches 1; Indels 42; Gaps 1;  
QY 1 catgcaataagagctctcacctgtctctctctgtgaccttgcgattcgtcttagtcg 60  
DB 71 catgcaataagagctctcacctgtctctctctgtgaccttgcgattcgtcttagtcg 130  
QY 61 caccagaagatactactggtggtcagtgcaactgtcatggaactatgcaaaagtgatct 120  
DB 131 caccagaagatactactggtggtcagtgcaactgtcatggaactatgcaaaagtgatct 190  
QY 121 cggggaagctgctgtgagcgaagaatttcctcctagagtgccaaaattcttccatcaa 180  
DB 191 cggggaagctgctgtgagcgaagaatttcctcctagagtgccaaaattcttccatcaa 250  
QY 181 caccctagctcggtgacaaaagactctgttctgtagaattccacggtccattcaaat 240  
DB 251 caccctagctcggtgacaaaagactctgttctgtagaattccacggtccattcaaat 310  
QY 241 cggtaagcgaagccaccctgagatggtctgctaggtccatccacccagctgaggttca 300  
DB 311 cggtaagcgaagccaccctgagatggtctgctaggtccatccacccagctgaggttca 370  
QY 301 tgataagctgtgctatcaacttaagaataagcttccatccctgcaacttcaagctgt 360  
DB 371 tgataagctgtgctatcaacttaagaataagcttccatccctgcaacttcaagctgt 430

QY 361 tgggtatccctacttgaaaagctctctgaaggagctgaaatgatgatcagaccagtc 420  
DB 431 tgggtatccctacttgaaaagctctctgaaggagctgaaatgatgatcagaccagtc 490  
QY 421 ggaagaagaagatgataaagctctccctgtgtggaagccatacatatgtctgcaggtcct 480  
DB 491 ggaagaagaagatgataaagctctccctgtgtggaagccatacatatgtctgcaggtcct 550  
QY 481 gaaagaagaatgctcaatgagctctgaccacactgtgcttaaccatcatcttctca 540  
DB 551 gaaagaagaatgctcaatgagctctgaccacactgtgcttaaccatcatcttctca 610  
QY 541 tgtggaacttggtlaaaagacttgaaatccaggtccatattggaagccctactagatgaga 600  
DB 611 tgtggaacttggtlaaaagacttgaaatccaggtccatattggaagccctactagatgaga 670  
QY 601 agggagcttgcccaagaaaagacacagacctgacacaaattatctactcttctgt 660  
DB 671 agggagcttgcccaagaaaagacacagacctgacacaaattatctactcttctgt 730  
QY 661 attgataagggaaaagattgcaactcagaacaaagaactcctgtatgcaggaatgaga 720  
DB 731 attgataagggaaaagattgcaactcagaacaaagaactcctgtatgcaggaatgaga 790  
QY 721 tgtgtcatctgtcgggcttggtcctaaatgacacagctcaatggtatglaaacaggtc 780  
DB 791 tgtgtcatctgtcgggcttggtcctaaatgacacagctcaatggtatglaaacaggtc 850  
QY 781 tctgcaggtctgattgtagtgcacaggaataatcagttctatgagcatgtatggaatggg 840  
DB 851 tctgcaggtctgattgtagtgcacaggaataatcagttctatgagcatgtatggaatggg 910  
QY 841 caccctcttgaagtgagcctcaatatctctcgaagtgacacacattcttcttgaggaaaca 900  
DB 911 caccctcttgaagtgagcctcaatatctctcgaagtgacacacattcttcttgaggaaaca 970  
QY 901 tcgcaagcgctcctctggaatctctgcacataacttctctactgctcaaacactctgtat 960  
DB 971 tcgcaagcgctcctctggaatctctgcacataacttctctactgctcaaacactctgtat 1030  
QY 961 ggaacttggaagatttctactcgttttgcataatctctccacacaaatgagtgcatagga 1020  
DB 1031 ggaacttggaagatttctactcgttttgcataatctctccacacaaatgagtgcatagga 1090  
QY 1021 agcttatgtcaagatgagcagctgtccagaggaaccccaactcagaatgaaaaataatga 1080  
DB 1091 agcttatgtcaagatgagcagctgtccagaggaaccccaactcagaatgaaaaataatga 1150  
QY 1081 agaagcgaagaactatgatagtactctactgattctgaaatgatagtgtgcaggttga 1140  
DB 1151 agaagcgaagaactatgatagtactctactgattctgaaatgatagtgtgcaggttga 1210  
QY 1141 tgatgacactctcctctcttaccataatcgttcaggttgccagaagaagatccataaac 1200  
DB 1211 tgatgacactctcctctcttaccataatcgttcaggttgccagaagaagatccataaac 1270  
QY 1201 ttgggtacattacattgtgtctgtagaagagagactgtggaactatctccctagtcctgc 1260  
DB 1271 ttgggtacattacattgtgtctgtagaagagagactgtggaactatctccctagtcctgc 1330  
QY 1261 ccccgatgacagaaggttaataaaatcaatattgaaacatgacctcagcaggtatgtgtag 1320  
DB 1331 ccccgatgacagaaggttaataaaatcaatattgaaacatgacctcagcaggtatgtgtag 1390  
QY 1321 gaagtacaaaaaagtcgattatgtagcatacacagatgaaacctttagagctcgtgtagc 1380  
DB 1391 gaagtacaaaaaagtcgattatgtagcatacacagatgaaacctttagagctcgtgtagc 1450  
QY 1381 tattcagatgaatacagaactcttggaacttcttactttaggggaattttagaacacact 1440  
DB 1451 tattcagatgaatacagaactcttggaacttcttactttaggggaattttagaacacact 1510



QY 1441 gtgtattatatttaagaatcaagcaagcaagcatalaacaattacccacggaatcac 1500  
|||||  
Db 1511 gtgtattatatttaagaatcaagcaagcatalaacaattacccacggaatcac 1570  
|||||  
QY 1501 tgaatccgttccttctgtatcttaagaagattaccaaaggtgttaaacatttgaagattc 1560  
|||||  
Db 1571 tgaatccgttccttctgtatcttaagaagattaccaaaggtgttaaacatttgaagattc 1630  
|||||  
QY 1561 tccaatcttcgcagaggaanaatattcaaatataaagtgaagcttgaagaattgggcc 1620  
|||||  
Db 1631 tccaatcttcgcagaggaanaatattcaaatataaagtgaagcttgaagaattgggcc 1690  
|||||  
QY 1621 aactaaatcagatccctcgtgtgcctgcagccgctattactcttgcttctgtaataagagag 1680  
|||||  
Db 1691 aactaaatcagatccctcgtgtgcctgcagccgctattactcttgcttctgtaataagagag 1750  
|||||  
QY 1681 agatctagcttcaggaactcaattggccctctccatctgcttaacaagaatctgtatgca 1740  
|||||  
Db 1751 agatctagcttcaggaactcaattggccctctccatctgcttaacaagaatctgtatgca 1810  
|||||  
QY 1741 aagaggaacacagataatgttcagacaagagaatgcatcctgttctctgtattgtga 1800  
|||||  
Db 1811 aagaggaacacagataatgttcagacaagagaatgcatcctgttctctgtattgtga 1870  
|||||  
QY 1801 gaaccggaagctgtgtaacctcaacagaaatataacaagcttctccccaatccagcttgagt 1860  
|||||  
Db 1871 gaaccggaagctgtgtaacctcaacagaaatataacaagcttctccccaatccagcttgagt 1930  
|||||  
QY 1861 gcagcttgaggaatccagagttccaaagcctccaaacatctgaacagcaatcgaattgata 1920  
|||||  
Db 1931 gcagcttgaggaatccagagttccaaagcctccaaacatctgaacagcaatcgaattgata 1990  
|||||  
QY 1921 ttgtgataagtttgcaagtttgcagagttgtttgtgataaggtgagcaatcgtgtaaccttaag 1980  
|||||  
Db 1991 ttgtgataagtttgcaagtttgcagagttgtttgtgataaggtgagcaatcgtgtaaccttaag 2050  
|||||  
QY 1981 catctgagcacaagactgaactcctctctgtctctctctctctctctctctctctctctctca 2040  
|||||  
Db 2051 catctgagcacaagactgaactcctctctgtctctctctctctctctctctctctctctctca 2110  
|||||  
QY 2041 aatgtctataagaacaacaactacacctatcccaattccagagagaacgtctctcatgtc 2100  
|||||  
Db 2111 aatgtctataagaacaacaactacacctatcccaattccagagagaacgtctctcatgtc 2170  
|||||  
QY 2101 gatggaacacccaggtctatgagattctgggtggtgcacaactcagacttttggagaacag 2160  
|||||  
Db 2171 gatggaacacccaggtctatgagattctgggtggtgcacaactcagacttttggagaacag 2230  
|||||  
QY 2161 catgaccgcttactgaagtttctgaatttgtaacaagaacactgtgtattatacagagga 2220  
|||||  
Db 2231 catgaccgcttactgaagtttctgaatttgtaacaagaacactgtgtattatacagagga 2290  
|||||  
QY 2221 cagttatgaagataattcagacatactgtctgagttaaaaacaatgtcaatlgaaaccaagaag 2280  
|||||  
Db 2291 cagttatgaagataattcagacatactgtctgagttaaaaacaatgtcaatlgaaaccaagaag 2350  
|||||  
QY 2281 ctctccccaagaattccaagacacctagcactagcagaagaacatttatgtccaccacc 2340  
|||||  
Db 2351 ctctccccaaga-----ccacc 2368  
|||||  
QY 2341 agtcttgaacgcacatcaacgcggaataactcgtactactcttcagtcagatacagaagga 2400  
|||||  
Db 2369 agtcttgaacgcacatcaacgcggaataactcgtactactcttcagtcagatacagaagga 2428  
|||||  
QY 2401 aatgtacatgataatccatatcagttgtaaatgaaagaagaagatttgaacatttga 2460  
|||||  
Db 2429 aatgtacatgataatccatatcagttgtaaatgaaagaagaagatttgaacatttga 2488  
|||||  
QY 2461 tgaagatgaaataatcagaagcccccgcagcttccaagaagaacacagacattttatgtc 2520  
|||||  
Db 2489 tgaagatgaaataatcagaagcccccgcagcttccaagaagaacacagacattttatgtc 2548  
|||||  
QY 2521 tgcagtgagaaggtctcgggattatgggatagagtagtccccaatgttcttaagaacag 2580  
|||||

Db 2549 tgcagtgagaaggtctcgggattatgggatagagtagtccccaatgttcttaagaacag 2608  
|||||  
QY 2581 ggtccagaagtggcaggtgtccctcagttccaagaagaagttgtttccagaatttactgattg 2640  
|||||  
Db 2609 ggtccagaagtggcaggtgtccctcagttccaagaagaagttgtttccagaatttactgattg 2668  
|||||  
QY 2641 ctcccttaccagcccttataccgttgagagactaaatgaacatttggactccctggggcc 2700  
|||||  
Db 2669 ctcccttaccagcccttataccgttgagagactaaatgaacatttggactccctggggcc 2728  
|||||  
QY 2701 atataagaagcagaagtgtgaagaataatcatgttgaacttccagaataatcagccctcgt 2760  
|||||  
Db 2729 atataagaagcagaagtgtgaagaataatcatgttgaacttccagaataatcagccctcgt 2788  
|||||  
QY 2761 tccctattcctcttatactcagccttattcttcatgagaagatcagaagcagaagcaga 2820  
|||||  
Db 2789 tccctattcctcttatactcagccttattcttcatgagaagatcagaagcagaagcaga 2848  
|||||  
QY 2821 accatagaanaaactgttcaagccataatgaacccaactactttggagaagtgcaca 2880  
|||||  
Db 2849 accatagaanaaactgttcaagccataatgaacccaactactttggagaagtgcaca 2908  
|||||  
QY 2881 tcatatgcccacataaagatgagttgactgcagaagcctgggcttatctctgtatgt 2940  
|||||  
Db 2909 tcatatgcccacataaagatgagttgactgcagaagcctgggcttatctctgtatgt 2968  
|||||  
QY 2941 tgaactggaanaaagatgttcgaactcagagcctgtatgtgaacccctctgtctgcacactaa 3000  
|||||  
Db 2969 tgaactggaanaaagatgttcgaactcagagcctgtatgtgaacccctctgtctgcacactaa 3028  
|||||  
QY 3001 cacactgaacccctgtcatatggagacaagtgacgtatcaagaatttctgttttcac 3060  
|||||  
Db 3029 cacactgaacccctgtcatatggagacaagtgacgtatcaagaatttctgttttcac 3088  
|||||  
QY 3061 catcttctgataagacccaagaagctgttacttcaactgaaataatgtgaagaagaactgcagggc 3120  
|||||  
Db 3089 catcttctgataagacccaagaagctgttacttcaactgaaataatgtgaagaagaactgcagggc 3148  
|||||  
QY 3121 tccctgcataatccagatgtgaagaatcccaacttttaagaagaatatactgcctccatgc 3180  
|||||  
Db 3149 tccctgcataatccagatgtgaagaatcccaacttttaagaagaatatactgcctccatgc 3208  
|||||  
QY 3181 caatggtctacataatgtgaatacactacccctgtgtgaatgtgcataagatccaaagattcg 3240  
|||||  
Db 3209 caatggtctacataatgtgaatacactacccctgtgtgaatgtgcataagatccaaagattcg 3268  
|||||  
QY 3241 atggtatctgtcagcattggcagcaatgaaacacatccattctatccattcagttgaca 3300  
|||||  
Db 3269 atggtatctgtcagcattggcagcaatgaaacacatccattctatccattcagttgaca 3328  
|||||  
QY 3301 tgtgttcaactgtacgaaaaaagaggagataaaatgagcactgtacacatctctacag 3360  
|||||  
Db 3329 tgtgttcaactgtacgaaaaaagaggagataaaatgagcactgtacacatctctacag 3388  
|||||  
QY 3361 tgttttgaagaacgttggaataatgttacatccaagcgtggaatttgggggttggaattgct 3420  
|||||  
Db 3389 tgttttgaagaacgttggaataatgttacatccaagcgtggaatttgggggttggaattgct 3448  
|||||  
QY 3421 tatgtgcgaacatctacatctcgtgagatgacacactttctcgtgtgtaacgcaataagtg 3480  
|||||  
Db 3449 tatgtgcgaacatctacatctcgtgagatgacacactttctcgtgtgtaacgcaataagtg 3508  
|||||  
QY 3481 tcaagactccctcgtggaatgtgctctcagacacatagaagatttccaagattcagcttcag 3540  
|||||  
Db 3509 tcaagactccctcgtggaatgtgctctcagacacatagaagatttccaagattcagcttcag 3568  
|||||  
QY 3541 acataatgagcagttgggcccacaagcgtgagcagactcatattccggatccaatcaatgc 3600  
|||||  
Db 3569 acataatgagcagttgggcccacaagcgtgagcagactcatattccggatccaatcaatgc 3628  
|||||  
QY 3601 ctggaagaccaagaagcccttctctgatacaaggtgagctgttgagacacatgattat 3660  
|||||



```

Db 3629 ctgagacccaaggagcccttcttgatcaagtgatctgtggaaccaatgattat 3688
Qy 3661 tcaagcga tcaagacagaggtgccccgcgcagaagttctccaagcctctacatctcaagt 3720
Db 3689 tcaagcga tcaagacagagaggggccccgcgcagaagttctccaagcctctacatctcaagt 3748
Qy 3721 tatcatcatgatatagctcttgatgggaagaagtgcagactctcgaaggaatctcaatg 3780
Db 3749 tatcatcatgatatagctcttgatgggaagaagtgcagactctcgaaggaatctcaatg 3808
Qy 3781 aaccttaatgtcttctcttggaatgatgatcatctgggataaacaacatattttaa 3840
Db 3809 aaccttaatgtcttctcttggaatgatgatcatctgggataaacaacatattttaa 3868
Qy 3841 cccccaatctctgcgcatacatcgtctgcaccaactctatagatctgcagcac 3900
Db 3869 cccccaatctctgcgcatacatcgtctgcaccaactctatagatctgcagcac 3928
Qy 3901 tcttcgaatggagttgagtgagcgtgtgatttaaatagttgcagcatgccaatgggaatgga 3960
Db 3929 tcttcgaatggagttgagtgagcgtgtgatttaaatagttgcagcatgccaatgggaatgga 3988
Qy 3961 gagtaagaacatatacagatgcacagatattacgtcttaccatcttaccataatgttgc 4020
Db 3989 gagtaagaacatatacagatgcacagatattacgtcttaccatcttaccataatgttgc 4048
Qy 4021 caccgtgtctctcccaaaagctcgactcacctcccaaggaggaagtaatgctctgagacc 4080
Db 4049 caccgtgtctctcccaaaagctcgactcacctcccaaggaggaagtaatgctctgagacc 4108
Qy 4081 tcaagtgtaataatccaaagaagtgcgcagaatggaatccagaagaacaatgaaagtcac 4140
Db 4109 tcaagtgtaataatccaaagaagtgcgcagaatggaatccagaagaacaatgaaagtcac 4168
Qy 4141 agggagtaactactcaaggagtaaaatctctgtcttaccaggaatgtaaggaagttcct 4200
Db 4169 agggagtaactactcaaggagtaaaatctctgtcttaccaggaatgtaaggaagttcct 4228
Qy 4201 catctccagcagtcacaagtgcacatcagtcgactcttcttcagaaatggaagtaaa 4260
Db 4229 catctccagcagtcacaagtgcacatcagtcgactcttcttcagaaatggaagtaaa 4288
Qy 4261 ggttttccaggaataatcaagactccttcaacactgtgtgaactcttcagaccacggt 4320
Db 4289 ggttttccaggaataatcaagactccttcaacactgtgtgaactcttcagaccacggt 4348
Qy 4321 actgaactcgactacttcaaatctcaacccccagagttggtgcacagatctgcctgaagat 4380
Db 4349 actgaactcgactacttcaaatctcaacccccagagttggtgcacagatctgcctgaagat 4408
Qy 4381 ggaaggtctgggctgcagagcagaggaactctactga 4417
Db 4409 ggaaggtctgggctgcagagcagaggaactctactga 4445

```

```

RESULT 6
AAV15338
ID AAV15338 standard; DNA: 4832 BP.
AC AAV15338:
XX
DT 20-JUL-1998 (first entry)
XX
DE Human Factor VIII SQN deletion mutant DNA.
XX
KM Factor VIII: blood clotting; haemophilia A; gene therapy.
XX
KW retrovirus; vector; human; ss.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT CDS 72..4445

```

```

FT XX /*tag= a
PN XX MO3800542-A2.
PD XX 08-JAN-1998.
XX
PF XX 02-JUL-1997; 97WO-US11785.
XX
PR XX 04-JUN-1997; 97US-0869309.
PR XX 03-JUL-1996; 96US-0645601.
PR XX 13-AUG-1996; 96US-0696381.
XX
PA (CHIR ) CHIRON CORP.
XX
PI Allen JR, Barber JR, Boder M, Chang SM, Chong K;
PI De la Vega D, Depolo NJ, Greengard J, Hsu DC, Ibanez CE;
PI Jolly DJ, Mittelstaedt DM, Prussak CE, Respass JG;
XX
XX WP1: 1998-086967/08.
DR P-PSDB: AAM44372.
XX
XX
PT New replication defective recombinant retroviruses - which express B
PT domain-deleted human factor VIII or human factor IX for the
PT treatment of haemophilia
XX
PS Claim 6; Page 174-175; 236pp; English.
XX
XX
CC This DNA sequence includes a coding region for the B domain
CC deletion mutant SQN (see AAM44372) of human Factor VIII. The SQN
CC mutant is created by fusing Ser-743 to Gln-1638 of native Factor
CC VIII (see AAM44373) to form a Ser-Gln-Asn (SQN) link between the A2
CC and A3 Factor VIII domains. When compared to plasmid-derived
CC Factor VIII, the SQN deletion does not influence the in vivo
CC pharmacokinetics, but the reduced size of the molecule appears to
CC decrease proteolytic degradation. The invention relates to
CC preparations of replication defective recombinant retrovirus (RV)
CC expressing a B domain-deleted human Factor VIII protein, where the
CC recombinant RV is capable of infecting human cells, is resistant
CC to degradation by human complement and is capable of inducing
CC long-term (at least 30 days and up to 6 months or longer
CC post-injection) systemic expression of Factor VIII when
CC administered to a haemophilia A patient.
XX
SQ Sequence 4832 BP; 1381 A; 1089 C; 1055 G; 1307 T; 0 other;

```

```

Query Match 97.8%; Score 4321.4; DB 19; Length 4832;
Best Local Similarity 99.0%; Pred. NO. 0; Mismatches 1; Indels 42; Gaps 1;
Matches 4374; Conservative 0;
Qy 1 catgcaaatagagctctccactgtcttctctgtctgtgctcttgagattcgtcttagtc 60
Db 71 catgcaaatagagctctccactgtcttctctgtctgtgctcttgagattcgtcttagtc 130
Qy 61 caccagaagatactactctggtgagtcagtcgaactgcatgaggactatataatgcaagtgcct 120
Db 131 caccagaagatactactctggtgagtcagtcgaactgcatgaggactatataatgcaagtgcct 190
Qy 121 cgggagctgctgtgagcgaagaattctcctcagaagtgccaataatctttcatcaca 180
Db 191 cgggagctgctgtgagcgaagaattctcctcagaagtgccaataatctttcatcaca 250
Qy 181 caccatactgtgatacaaaaagactctgtgtgataatcaacggttcaacttcaacat 240
Db 251 caccatactgtgatacaaaaagactctgtgtgataatcaacggttcaacttcaacat 310
Qy 241 cgttaagcgaagccaccctgagatggtctgctagatgctcctacacatcagaagtcaggttta 300
Db 311 cgttaagcgaagccaccctgagatggtctgctagatgctcctacacatcagaagtcaggttta 370
Qy 301 tgatacagtggtcatttaacttaagaacaatggtcttccatccctgtcagttcattgctgt 360
Db 371 tgatacagtggtcatttaacttaagaacaatggtcttccatccctgtcagttcattgctgt 430

```



QY 361 tgggtatccactggaagcttcctgaggagctgaaatgatgatccagacagtc aaag 420  
|||||  
Db 431 tgggtatccactggaagcttcctgaggagctgaaatgatgatccagacagtc aaag 480  
QY 421 gggagaaagatgataaagtcctccctggcggaaagcatatagtcgcgagtcctc 480  
|||||  
Db 491 gggagaaagatgataaagtcctccctggcggaaagcatatagtcgcgagtcctc 550  
QY 481 gaaagaaatggtccaaatggtccttgacccactggtgccttaactacatactctc 540  
|||||  
Db 551 gaaagaaatggtccaaatggtccttgacccactggtgccttaactacatactctc 610  
QY 541 tgggaaacctgtaaaagacttgaaatcaggcctcattggagccctactagtcgtaga 600  
|||||  
Db 611 tgggaaacctgtaaaagacttgaaatcaggcctcattggagccctactagtcgtaga 670  
QY 601 agggagcttgccaaaggaaagacacagaccttgacacaaattataactattctgtc 660  
|||||  
Db 671 agggagcttgccaaaggaaagacacagaccttgacacaaattataactattctgtc 730  
QY 661 atttgatgaagggaagatggtgcacccagaaacaaagaactccttgatcgagatgga 720  
|||||  
Db 731 atttgatgaagggaagatggtgcacccagaaacaaagaactccttgatcgagatgga 790  
QY 721 tggctgacatcgctccggcgcctgacaaatgacacagtcacatggttatgttaaagc 780  
|||||  
Db 791 tggctgacatcgctccggcgcctgacaaatgacacagtcacatggttatgttaaagc 850  
QY 781 tctgcagagctgctgattggaatgacacaggaatcagcttatgtgacatggtgaatg 840  
|||||  
Db 851 tctgcagagctgctgattggaatgacacaggaatcagcttatgtgacatggtgaatg 910  
QY 841 caccaactcctgaaagtgcactccaatatctcctcgaaagtcacacattcctgtgagga 900  
|||||  
Db 911 caccaactcctgaaagtgcactccaatatctcctcgaaagtcacacattcctgtgagga 970  
QY 901 tcccgagcgctccttggaatctcgcgcaaatcttccttaactgctgcacaaactctgat 960  
|||||  
Db 971 tcccgagcgctccttggaatctcgcgcaaatcttccttaactgctgcacaaactctgat 1030  
QY 961 ggaacctggaagcttctactggttctgtcatatctctccacacaaatgcatgga 1020  
|||||  
Db 1031 ggaacctggaagcttctactggttctgtcatatctctccacacaaatgcatgga 1090  
QY 1021 agcttatgctaaagttagacagctgctcagaggaaccccaactcagatgaaataatga 1080  
|||||  
Db 1091 agcttatgctaaagttagacagctgctcagaggaaccccaactcagatgaaataatga 1150  
QY 1081 agaaagcgaagactatgatatgacttactgattctgtaaatgagtggtgcagtttga 1140  
|||||  
Db 1151 agaaagcgaagactatgatatgacttactgattctgtaaatgagtggtgcagtttga 1210  
QY 1141 tgaatgaactctcctcctcttataccaatctgcctcagttgccaagaagatccta aac 1200  
|||||  
Db 1211 tgaatgaactctcctcctcttataccaatctgcctcagttgccaagaagatccta aac 1270  
QY 1201 ttgggttaactatattgctgctgtaagaagagagctgggactatgctccctta gctcctgc 1260  
|||||  
Db 1271 ttgggttaactatattgctgctgtaagaagagagctgggactatgctccctta gctcctgc 1330  
QY 1261 ccccgatgacagaagctatataaagctcaatatctgaaacaatggtccctcagagtggt 1320  
|||||  
Db 1331 ccccgatgacagaagctatataaagctcaatatctgaaacaatggtccctcagagtggt 1390  
QY 1321 gaaatgaacaaagtcogatttatggtacatacacagatgaaacctttaagactcgt 1380  
|||||  
Db 1391 gaaatgaacaaagtcogatttatggtacatacacagatgaaacctttaagactcgt 1450  
QY 1381 tattcagcatgaatcaggaatctctggaccttaactttatggggaagttggaaacact 1440  
|||||  
Db 1451 tattcagcatgaatcaggaatctctggaccttaactttatggggaagttggaaacact 1510

QY 1441 gttcatatatttaagaatcaagcaagcaccataataacatctaccctcaggaatcac 1500  
|||||  
Db 1511 gttcatatatttaagaatcaagcaagcaccataataacatctaccctcaggaatcac 1570  
QY 1501 tgaatgcgtcctcttgatccaaggagatataccaagaagtgtaaaaacatttgaagatt 1560  
|||||  
Db 1571 tgaatgcgtcctcttgatccaaggagatataccaagaagtgtaaaaacatttgaagatt 1630  
QY 1561 tccaattcgcagaggaataatccaataataatggaacgtgactgtagaatgagcc 1620  
|||||  
Db 1631 tccaattcgcagaggaataatccaataataatggaacgtgactgtagaatgagcc 1690  
QY 1621 aactaaatcagatccctcggtgccttgaccgcctatctactcagttcttgtaatatgagag 1680  
|||||  
Db 1691 aactaaatcagatccctcggtgccttgaccgcctatctactcagttcttgtaatatgagag 1750  
QY 1681 agatctagcttcaggaatcattggccctcctcactatgctgtaaaaagatcgtatgata 1740  
|||||  
Db 1751 agatctagcttcaggaatcattggccctcctcactatgctgtaaaaagatcgtatgata 1810  
QY 1741 aagaggaacacgaataatgtaacagaaggaatgtaacatcctgtttctgtattatga 1800  
|||||  
Db 1811 aagaggaacacgaataatgtaacagaaggaatgtaacatcctgtttctgtattatga 1870  
QY 1801 gaaacgaagctggttacctccaagaaatatacaacgcttctcccaatccagctgagat 1860  
|||||  
Db 1871 gaaacgaagctggttacctccaagaaatatacaacgcttctcccaatccagctgagat 1930  
QY 1861 gcaagcttgaaatccagagatgccaaagcctccaacatcatgacagatcagtcatagt 1920  
|||||  
Db 1931 gcaagcttgaaatccagagatgccaaagcctccaacatcatgacagatcagtcatagt 1990  
QY 1921 ttttgatagtttcagatggttcagttggttgcataggagtggtactggtatcttaag 1980  
|||||  
Db 1991 ttttgatagtttcagatggttcagttggttgcataggagtggtactggtatcttaag 2050  
QY 1981 catgtgagcacaagactgactcctctgtcctcctcctcctcgatataccttcaaacaa 2040  
|||||  
Db 2051 catgtgagcacaagactgactcctctgtcctcctcctcctcgatataccttcaaacaa 2110  
QY 2041 aatggtctatgaaagacacactccaactatcccatctcagaggaacatgcttctca 2100  
|||||  
Db 2111 aatggtctatgaaagacacactccaactatcccatctcagaggaacatgcttctca 2170  
QY 2101 gatggaacacccaggtctatggaattctgggtggtcccaactcagacttccggaacagag 2160  
|||||  
Db 2171 gatggaacacccaggtctatggaattctgggtggtcccaactcagacttccggaacagag 2230  
QY 2161 catgacgcgcttactgaaggttctctagtttgtaacagaacactggtgatatacagaga 2220  
|||||  
Db 2231 catgacgcgcttactgaaggttctctagtttgtaacagaacactggtgatatacagaga 2290  
QY 2221 cagttatgaagatatttcagcatactgctgtagttaaacaatgcatatgaaacaaag 2280  
|||||  
Db 2291 cagttatgaagatatttcagcatactgctgtagttaaacaatgcatatgaaacaaag 2350  
QY 2281 ctctccccaagaattccaagacacctgacactagggcaaaagcaatttaatgccaacccac 2340  
|||||  
Db 2351 ctctccccaagaattccaagacacctgacactagggcaaaagcaatttaatgccaacccac 2368  
QY 2341 agctctgaaacgcctcaacagcggaaataaactcgtactactcctcagtcagatccaagaga 2400  
|||||  
Db 2369 agctctgaaacgcctcaacagcggaaataaactcgtactactcctcagtcagatccaagaga 2428  
QY 2401 aattgactatgatacatatcagttgaatgaaggaaggaatttgaacttatga 2460  
|||||  
Db 2429 aattgactatgatacatatcagttgaatgaaggaaggaatttgaacttatga 2488  
QY 2461 tgaagatgaaatcagaagcccgagccttcaaaagaaacaagcactattttatg 2520  
|||||  
Db 2489 tgaagatgaaatcagaagcccgagccttcaaaagaaacaagcactattttatg 2548  
QY 2521 tgcagtggaagagctcctggtattatgagtgagtaagctccccaatgcttctaagaacag 2580



|||||  
Db 2549 tgcagctggaagagctccggagctta tggatctgtagctcccccacatgcttaagaacag 2608  
OY 2581 ggcctcaagctggcagctgctccctcagttcaagaagctgtttccaggaattactgacg 2640  
Db 2609 ggcctcaagctggcagctgctccctcagttcaagaagctgtttccaggaattactgacg 2668  
OY 2641 ctccttactcagcctctataccgctggagaactaaatgaacatttggagctctggggcc 2700  
Db 2669 ctccttactcagcctctataccgctggagaactaaatgaacatttggagctctggggcc 2728  
OY 2701 atatatagaagcagaagtgtgaagataataatactgttaacttccgaaatccagctctcg 2760  
Db 2729 atataagaagcagaagtgtgaagataataatactgttaacttccgaaatccagctctcg 2788  
OY 2761 tccctatctctctatctctagcctctatttctta tgaagaaatccaggaagcagagcaga 2820  
Db 2789 tccctatctctctatctctagcctctatttctta tgaagaaatccaggaagcagagcaga 2848  
OY 2821 acctagaaaaaacttctgcaagccttaa tgaaccaaaacttacttctggaagtgcaca 2880  
Db 2849 acctagaaaaaacttctgcaagccttaa tgaaccaaaacttacttctggaagtgcaca 2908  
OY 2881 tcatatggcccccaataagaatgagttgactgtcaagcctgggcttaattctctgagt 2940  
Db 2909 tcatatggcccccaataagaatgagttgactgtcaagcctgggcttaattctctgagt 2968  
OY 2941 tgcacctggaanaagatgtgcactcagccttgatggacccttctggtctgcacactaa 3000  
Db 2969 tgcacctggaanaagatgtgcactcagccttgatggacccttctggtctgcacactaa 3028  
OY 3001 caactagaccctgtctcatctggagacaa tgaagctacagtaagaattctctgttttcac 3060  
Db 3029 caactagaccctgtctcatctggagacaa tgaagctacagtaagaattctctgttttcac 3088  
OY 3061 catctcttgatgagaccaaaagctgttacttca ctgaaaaataatggaagaagaactgcagggc 3120  
Db 3089 catctcttgatgagaccaaaagctgttacttca ctgaaaaataatggaagaagaactgcagggc 3148  
OY 3121 tccctgcaatccaga tggagaatcccaacttcaagaagatatactctccctgcacat 3160  
Db 3149 tccctgcaatccaga tggagaatcccaacttcaagaagatatactctccctgcacat 3208  
OY 3181 caatggtctacataatgatacactactgtgct tagtaatggtcgaagatacaagaatctg 3240  
Db 3209 caatggtctacataatgatacactactgtgct tagtaatggtcgaagatacaagaatctg 3268  
OY 3241 atggtatctgtcagccttggcagcaatgaa aacatccatcttaatttcaattgagca 3300  
Db 3269 atggtatctgtcagccttggcagcaatgaa aacatccatcttaatttcaattgagca 3328  
OY 3301 tggcttcaactgatacgaanaaaagaggag taaataatgacactgtacaaatctctaccag 3360  
Db 3329 tggcttcaactgatacgaanaaaagaggag taaataatgacactgtacaaatctctaccag 3388  
OY 3361 tggcttctgagacagctggaatgttaccatcca aagctggaatttggcggtggaatgct 3420  
Db 3389 tggcttctgagacagctggaatgttaccatcca aagctggaatttggcggtggaatgct 3448  
OY 3421 tatctggcagcatctaaatgtctggatgagcaac cttttctgtgtgtacacataaagt 3480  
Db 3449 tatctggcagcatctaaatgtctggatgagcaac cttttctgtgtgtacacataaagt 3508  
OY 3481 tcaagactcccttggaaatgcttctgacacataga gatattcagatttaagcttcagg 3540  
Db 3509 tcaagactcccttggaaatgcttctgacacataga gatattcagatttaagcttcagg 3568  
OY 3541 acaata tgaacagctggccccaaaagctggtccaga cctcatatctcggatcaatcaatgct 3600  
Db 3569 acaata tgaacagctggccccaaaagctggtccaga cctcatatctcggatcaatcaatgct 3628  
OY 3601 ctggagcaccaaagagccctttcttgatcaagctg gatctgtgtgaccaaagtata 3660  
|||||

Db 3629 ctggagcaccaaagagcccttcttgatcaagg tgaatctgttggcaccaatgatat 3688  
OY 3661 tcaagcga tcaagaccagggctgccctcagaagtctcc agcctctacatctcaagt 3720  
Db 3689 tcaagcga tcaagaccagggctgccctcagaagtctcc agcctctacatctcaagt 3748  
OY 3721 tatcatatgataagcttga tgggaagaagtggcaga cttatccgaatccactg 3780  
Db 3749 tatcatatgataagcttga tgggaagaagtggcaga cttatccgaatccactg 3808  
OY 3781 aaccttaatggtctcttcttgcaatgtggatca tctcgggaataaacaatatctttta 3840  
Db 3809 aaccttaatggtctcttcttgcaatgtggatca tctcgggaataaacaatatctttta 3868  
OY 3841 cccctcaatctatctcgatca tccgcttggaccacca tcatataga ttcgcagcac 3900  
Db 3869 cccctcaatctatctcgatca tccgcttggaccacca tcatataga ttcgcagcac 3928  
OY 3901 tcttcgcagtgaatgtga tgggtgtgaatttaaa tagttgcagcatgccaatggaa tgg 3960  
Db 3929 tcttcgcagtgaatgtga tgggtgtgaatttaaa tagttgcagcatgccaatggaa tgg 3988  
OY 3961 gagtaagaacatatacaga tgcagatctactgctc acttcaacttcaacatatgtttgc 4020  
Db 3989 gagtaagaacatatacaga tgcagatctactgctc acttcaacttcaacatatgtttgc 4048  
OY 4021 caactgtctctcctcaaaagctcgaactcacc tcccaaggaggaatgaatcctggagac 4080  
Db 4049 caactgtctctcctcaaaagctcgaactcacc tcccaaggaggaatgaatcctggagac 4108  
OY 4081 tcaagtgaaataatccaaa agatgtgctgcagaatg gacttccagaagaacaaatgaagtca 4140  
Db 4109 tcaagtgaaataatccaaa agatgtgctgcagaatg gacttccagaagaacaaatgaagtca 4168  
OY 4141 aggaatataactatacga tgggaataatctctgcttaccagaatgatatgtgaaggagtctct 4200  
Db 4169 aggaatataactatacga tgggaataatctctgcttaccagaatgatatgtgaaggagtctct 4228  
OY 4201 catctcagcagctacaaga tggccatcagtgactctcttcttccaagaatgcaaaataaa 4260  
Db 4229 catctcagcagctacaaga tggccatcagtgactctcttcttccaagaatgcaaaataaa 4288  
OY 4261 ggttttccaagggaatac aaagactcctcaacac tctgtgtgaactctcagaccacggt 4320  
Db 4289 ggttttccaagggaatac aaagactcctcaacac tctgtgtgaactctcagaccacggt 4348  
OY 4321 actgaactcgactacctt ggaattcacccccagag tttgggtgcaccagaatgccccttggat 4380  
Db 4349 actgaactcgactacctt ggaattcacccccagag tttgggtgcaccagaatgccccttggat 4408  
OY 4381 ggaagttctgggtcgag gacaggaactctactga 4417  
Db 4409 ggaagttctgggtcgag gacaggaactctactga 4445  
|||||

RESULT 7  
AAx82259  
ID AAx82259 standard; cDNA: 9164 BP.  
XX  
AC AAx82259;  
XX  
DT 18-AUG-1999 (first entry)  
XX  
DE Beta-domain deleted Factor VIII protein encoding gene (uncorrected).  
XX  
KW Factor VIII protein; gene modification; gene therapy; clinical disorder;  
KW splicing pattern; RNA processing; gene regulation; beta-domain; human;  
XX  
OS Homo sapiens.  
XX  
PN W09929848-A1.  
XX



PD 17-JUN-1999.  
XX 25-NOV-1998; 98MO-US25354.  
XX 16-JAN-1998; 98US-0071596.  
PR 05-DEC-1997; 97US-0067614.  
XX  
PA (IMMUNO-) IMMUNE RESPONSE CORP.  
XX  
PI BldIngaier S, Gonzales JEN, III CR, Yang CQ;  
XX  
XX WPI, 1999-385602/32.  
DR P-PSDB; AAY21675.  
XX  
XX Genes and vectors exhibiting increased expression and novel splicing  
PT patterns, useful for expression of, e.g. beta-domain deleted factor  
VIII  
XX  
XX Example 2; Page 79-89; 123pp; English.  
PS  
XX The invention describes novel genes and vectors exhibiting increased  
CC expression and novel splicing patterns. It provides a gene encoding a  
CC Factor VIII protein, that comprises one or more consensus or near  
CC consensus splice sites which have been corrected to increase expression.  
CC The method, DNA sequences and expression vectors can be used to increase  
CC the expression of a gene, especially a Factor VIII gene. Genes containing  
CC modified 5' and/or 3' untranslated regions have optimized expression  
CC levels and tissue-specific expression. The methods are used for  
CC identification and correction of consensus splice sites, addition of  
CC introns, optimization of 5' and 3' untranslated regions and increase in  
CC cytoplasmic RNA accumulation. Hence the DNA are useful in gene therapy  
CC to treat a clinical disorder. To study RNA processing and/or gene  
CC regulation. The present sequence represents an uncorrected version of  
CC beta-domain deleted factor VIII protein encoding gene sequence (construct  
CC pCY-2). This was used to develop a new coding sequence for beta-domain  
CC deleted factor VIII protein by correcting the consensus splice sites.  
XX  
XX Sequence 9164 BP; 2451 A; 2170 C; 2081 G; 2462 T; 0 other;  
SQ

Query Match 97.8%; Score 4321.4; DB 20; Length 9164;  
Best Local Similarity 99.0%; Pred. No. 0;  
Matches 4374; Conservative 0; Mismatches 1; Indels 42; Gaps 1;

QY 1 catgcaaatagactctccactcctgtcttcttctgtgctcttgatctgctttagtc 60  
DB 1005 catggaataagagctccacccctgtcttcttctgtgctcttgcgattcgtcttagtc 1064

QY 61 caccagaatactactacctggtgagtgagtgactgtcatgtgactatatagtcaagtgtct 120  
DB 1065 caccagaagagctactactggtgagtgagtgactgtcatgtgactatatagtcaagtgtct 1124

QY 121 cggtagagctcgtgtgagcaagattctcctcctagaagtgccaaatctttccatcaa 180  
DB 1125 cggtagagctcgtgtgagcaagattctcctcctagaagtgccaaatctttccatcaa 1184

QY 181 cactcaagtgtgtacaaaaaagactcgtttgtagaattcaaggttcaacttttcaacat 240  
DB 1185 cactcaagtgtgtacaaaaaagactcgtttgtagaattcaaggttcaacttttcaacat 1244

QY 241 cgtctaaagcacaagccacctgtgagtggtctgtcgtcgtcgtcctacatccaggctgagttta 300  
DB 1245 cgtctaaagcacaagccacctgtgagtggtctgtcgtcgtcgtcctacatccaggctgagttta 1304

QY 301 tgaatacagtggttcaacttaagaacatgtgcttccatcctgtcagttcattatgctgt 360  
DB 1305 tgaatacagtggttcaacttaagaacatgtgcttccatcctgtcagttcattatgctgt 1364

QY 361 tggatgatacctactgtgaaagctcttgaggagtgatgatatgatataagccagtgtaag 420  
DB 1365 tggatgatacctactgtgaaagctcttgaggagtgatgatatgatataagccagtgtaag 1424

QY 421 ggagaaagaagatgataaagttcttccctgtgtgagaagcacaataatgtctgagagtcct 480

DB 1425 ggagaaagaagatgataaagttcttccctgtgtgagaagcacaataatgtctgagagtcct 1484  
QY 481 gaaagaagaatgttccaatgtgctcttgacccactgtgcttaccatccatcttctca 540  
DB 1485 gaaagaagaatgttccaatgtgctcttgacccactgtgcttaccatccatcttctca 1544  
QY 541 tgtgacctgtgtaaagacttgaaatcgaagctcatttggaagccctactagatgtaga 600  
DB 1545 tgtgacctgtgtaaagacttgaaatcgaagctcatttggaagccctactagatgtaga 1604  
QY 601 agggagctctggccaagaaaagacacagacctgtgcacaatttaactacttttgcgt 660  
DB 1605 agggagctctggccaagaaaagacacagacctgtgcacaatttaactacttttgcgt 1664  
QY 661 atttgatgaagggaaaagtttgacactcagaacaagaagacctcttgatgcagatagga 720  
DB 1665 atttgatgaagggaaaagtttgacactcagaacaagaagacctcttgatgcagatagga 1724  
QY 721 tgcgtcatctgtcgtggccttggtcctaaatgacacagatgaatgttatgtaaaaggtc 780  
DB 1725 tgcgtcatctgtcgtggccttggtcctaaatgacacagatgaatgttatgtaaaaggtc 1784  
QY 781 tctgcaggtctgattgtgattccacaagaaaatcagtcattgtgcatgtgattgtgaatgg 840  
DB 1785 tctgcaggtctgattgtgattccacaagaaaatcagtcattgtgcatgtgattgtgaatgg 1844  
QY 841 caccactctgaagtgactcacaatattcctcgaagtgatcaacattcttgtgaagaacca 900  
DB 1845 caccactctgaagtgactcacaatattcctcgaagtgatcaacattcttgtgaagaacca 1904  
QY 901 tcggcaaggtcctcttggaatcttcgcaataacttcttactgtcttaaaacatttgat 960  
DB 1905 tcggcaaggtcctcttggaatcttcgcaataacttcttactgtcttaaaacatttgat 1964  
QY 961 ggaacttgagcaagtttctactgttttgcataatcttccccaacaatgatagtgcatgta 1020  
DB 1965 ggaacttgagcaagtttctactgttttgcataatcttccccaacaatgatagtgcatgta 2024  
QY 1021 agcttatgtcaaatgtagacagctgttccagaggaaccccaactacgaatgaaaataatga 1080  
DB 2025 agcttatgtcaaatgtagacagctgttccagaggaaccccaactacgaatgaaaataatga 2084  
QY 1081 agaagcggaagactatgatagtacttactgattctcgaatgtgagtgtgcgaagtttga 1140  
DB 2085 agaagcggaagactatgatagtacttactgattctcgaatgtgagtgtgcgaagtttga 2144  
QY 1141 tgaatgaacactctcctcttcaatccaaatcgtctcagttgcagaagacatccctaaac 1200  
DB 2145 tgaatgaacactctcctcttcaatccaaatcgtctcagttgcagaagacatccctaaac 2204  
QY 1201 ttgggtacatatacatgtcgtctgagaagagagactggtgactatgtctcccttagtcctgc 1260  
DB 2205 ttgggtacatatacatgtcgtctgagaagagagactggtgactatgtctcccttagtcctgc 2264  
QY 1261 ccccgatgacagaagttataaagtcaatatgtgaacaatggccctaaaggagttgtag 1320  
DB 2265 ccccgatgacagaagttataaagtcaatatgtgaacaatggccctaaaggagttgtag 2324  
QY 1321 gaagtacaataaagttccgatttaatgacatacagaagaaacctttaagactctgtagag 1380  
DB 2325 gaagtacaataaagttccgatttaatgacatacagaagaaacctttaagactctgtagag 2384  
QY 1381 tattcagcatgtaacaggaacttgggaaccttaacttattgtgggaagtgtgagacacact 1440  
DB 2385 tattcagcatgtaacaggaacttgggaaccttlaacttattgtgggaagtgtgagacacact 2444  
QY 1441 gttgattatatttaagaataaagcaagcagacatataaactacacctaaggaatcac 1500  
DB 2445 gttgattatatttaagaataaagcaagcagacatataaactacacctaaggaatcac 2504  
QY 1501 tgaatgcgtcctcttgcattcacaaggaattacccaaggtgtaaaacatttgaaggaattc 1560



Db 2505 tgaatcgcctccttctgtatctcaaggagatactcaaaaggctgtaaaacatctgaagattc 2564  
OY 1561 tccaattctgcagagagaataattccaataataatggaagtgaactgtagaagatggcc 1620  
Db 2565 tccaattctgcagagagaataataattccaataataatggaagatgtagaagatggcc 2624  
OY 1621 aactaaatcagatcctcgtgctgcgaaccgctatbactctagtttcgttaataatgagag 1680  
Db 2625 aactaaatcagatcctcgtgctgcgaaccgctatbactctagtttcgttaataatgagag 2684  
OY 1681 agatctagttctgaactcatctgccccctctcctcatctcgtcacaaaagaaatcgtatgaca 1740  
Db 2685 agatctagttctgaactcatctgccccctctcctcatctcgtcacaaaagaaatcgtatgaca 2744  
OY 1741 aagaagaaacacagataatgttcagacaagaggaatgtcatccctgtttctgttatgttga 1800  
Db 2745 aagaagaaacacagataatgttcagacaagaggaatgtcatccctgtttctgttatgttga 2804  
OY 1801 gaaccgaagctgtagctcctcacagaagataatacaacgcttctccccaatccagctgagat 1860  
Db 2805 gaaccgaagctgtagctcctcacagaagataatacaacgcttctccccaatccagctgagat 2864  
OY 1861 gcaagcttgaagatcccaagatcccaagctcccaacatcatctgacagcatcaatgtgtatgt 1920  
Db 2865 gcaagcttgaagatcccaagatcccaagctcccaacatcatctgacagcatcaatgtgtatgt 2924  
OY 1921 ttctgtagtttgtagtctgtagtttgttctgcatgaggtgcatatcgttcatcttaag 1980  
Db 2925 ttctgtagtttgtagtctgtagtttgttctgcatgaggtgcatatcgttcatcttaag 2984  
OY 1981 catctgagacacagatgaactcctctctcgtctcctctctcgtgataataccttcaaacaca 2040  
Db 2985 catctgagacacagatgaactcctctctcgtctcctctctcgtgataataccttcaaacaca 3044  
OY 2041 aatggtctatgaagaacacactcacacctatctccatctcagagagaagaatgtcttatctc 2100  
Db 3045 aatggtctatgaagaacacacactcacacctatctccatctcagagagaagaatgtcttatctc 3104  
OY 2101 gatcgaaaacacacaggtctatgtagtcttgagggtgcacacacacagactctcgaaacagag 2160  
Db 3105 gatcgaaaacacacaggtctatgtagtcttgagggtgcacacacacagactctcgaaacagag 3164  
OY 2161 catgacgcgcttactgtaggtttcttagttgtgacagaagaacacgtgtgatttaccagaga 2220  
Db 3165 catgacgcgcttactgtaggtttcttagttgtgacagaagaacacgtgtgatttaccagaga 3224  
OY 2221 caattagaagaataattcagatactgtgtgaatgaagaacatgcatctgtaacagagag 2280  
Db 3225 caattagaagaataattcagatactgtgtgaatgaagaacatgcatctgtaacagagag 3284  
OY 2281 ctctctcccaaatcagaagacacccctagcaactagcnaaagaatcttaatgcaacccacac 2340  
Db 3285 ctctctcccaaga-----ccacac 3302  
OY 2341 agctctgaaaacgcacatcaacgggaataaactcgtactactcttcaagtcagaatcaagaga 2400  
Db 3303 agctctgaaaacgcacatcaacgggaataaactcgtactactcttcaagtcagaatcaagaga 3362  
OY 2401 aattgactatgtagatcccatatcagtgtgaatgaagaagaagaatttgcacattatga 2460  
Db 3363 aattgactatgtagatcccatatcagtgtgaatgaagaagaagaatttgcacattatga 3422  
OY 2461 tggagatgaataatcaagcccccgagcttccaaaagaacaacacacttttatctgc 2520  
Db 3423 tggagatgaataatcaagcccccgagcttccaaaagaacaacacacttttatctgc 3482  
OY 2521 tgcagtgagaagcctctggaattatvggaatgagtagctcccaacatgcttcaagaacag 2580  
Db 3483 tgcagtgagaagcctctggaattatvggaatgagtagctcccaacatgcttcaagaacag 3542  
OY 2581 ggcctcaagtgagcagtgctccctcagttcaagaagaagtgtttccaaggaatttactcagtg 2640  
Db 3543 ggcctcaagtgagcagtgctccctcagttcaagaagaagtgtttccaaggaatttactcagtg 3602

OY 2641 ctcccttactcagcccttataccgttgagaactaaatgaacatttggagctcctgggccc 2700  
Db 3603 ctcccttactcagcccttataccgttgagaactaaatgaacatttggagctcctgggccc 3662  
OY 2701 atataagaagcagaagttgaagataataatggttaacttccaagaatcagcctctcg 2760  
Db 3663 atataagaagcagaagttgaagataataatggttaacttccaagaatcagcctctcg 3722  
OY 2761 tcccatcctcttatcttaagcttattcttgaagagaagataagaaggcaaggagaga 2820  
Db 3723 tcccatcctcttatcttaagcttattcttgaagagaagataagaaggcaaggagaga 3782  
OY 2821 aactagaaaaacttctgcaagcctaatagaacaaacttacttctgaaatgacaca 2880  
Db 3783 aactagaaaaacttctgcaagcctaatagaacaaacttacttctgaaatgacaca 3842  
OY 2881 tcaatgacacccacataagaatgagttgactgcaaaagcctgggtatctctgagt 2940  
Db 3843 tcaatgacacccacataagaatgagttgactgcaaaagcctgggtatctctgagt 3902  
OY 2941 tgaacttgaaaaaagatgtagcactcaggtctgattggagccctctgtgtgcacactaa 3000  
Db 3903 tgaacttgaaaaaagatgtagcactcaggtctgattggagccctctgtgtgcacactaa 3062  
OY 3001 cacactgaaccctgtcatctgggagacaagtgaacagtaacagaatttgcctgttttcac 3060  
Db 3963 cacactgaaccctgtcatctgggagacaagtgaacagtaacagaatttgcctgttttcac 4022  
OY 3061 catcttgaatgagaccaaagctgttacttcaactggaataataatggaagaacacgtcagagc 3120  
Db 4023 catcttgaatgagaccaaagctgttacttcaactggaataataatggaagaacacgtcagagc 4082  
OY 3121 tccctgcaatctcagatggaagatcccaactttaagaagaattatgccttccatgcaat 3180  
Db 4083 tccctgcaatctcagatggaagatcccaactttaagaagaattatgccttccatgcaat 4142  
OY 3181 caatgctacataatgatactacactcctggttagtaatggtctcagagatcaaaagattcg 3240  
Db 4143 caatgctacataatgatactacactcctggttagtaatggtctcagagatcaaaagattcg 4202  
OY 3241 atggtatctgctcagcagtggtgagcagaatgaagaacatccatctctatcttcaatgtagaca 3300  
Db 4203 atggtatctgctcagcagtggtgagcagaatgaagaacatccatctctatcttcaatgtagaca 4262  
OY 3301 tgtgtcactgttacgnaaaaagaggagtataaaatggcactgtgacatcctatccag 3360  
Db 4263 tgtgtcactgttacgnaaaaagaggagtataaaatggcactgtgacatcctatccag 4322  
OY 3361 tgtttttgagacagtgaatggtttaccatcccaagcttgnaatttggtggtggaatgct 3420  
Db 4323 tgtttttgagacagtgaatggtttaccatcccaagcttgnaatttggtggtggaatgct 4382  
OY 3421 tatggtcgagcatctacatctgctggatgtagacacacttttctgtgtgacagcaatagtg 3480  
Db 4383 tatggtcgagcatctacatctgctggatgtagacacacttttctgtgtgacagcaatagtg 4442  
OY 3481 tcaagctccctctgggaatggtctctgacacatatgaagatttcaagtttcaagttcag 3540  
Db 4443 tcaagctccctctgggaatggtctctgacacatatgaagatttcaagtttcaagttcag 4502  
OY 3541 acaatfatgagacatggtggcccaaaagcttggtccagactcattatctcgatcaatcaatg 3600  
Db 4503 acaatfatgagacatggtggcccaaaagcttggtccagactcattatctcgatcaatcaatg 4562  
OY 3601 ctggaagcaacagaagcccttctcttgatcaagtgagatctgtgtgacacaaatgatat 3660  
Db 4563 ctggaagcaacagaagcccttctcttgatcaagtgagatctgtgtgacacaaatgatat 4622  
OY 3661 tcaaggtacataagaacccaggtgtcccggtcagaagttctccagcctctacatctcagtt 3720  
Db 4623 tcaaggtacataagaacccaggtgtcccggtcagaagttctccagcctctacatctcagtt 4682



QY 3721 tatcatcatgtatagttctgtatggaagaatggtgacagattatcagaaattccactg 3780  
|||||  
Db 4683 tatcatcatgtatagttctgtatggaagaatggtgacagattatcagaaattccactg 4742  
QY 3781 aaccttaatggtctctcttcttggaatgtgtatcattctggtgataaacaataattttaa 3840  
|||||  
Db 4743 aaccttaatggtctctcttcttggaatgtgtatcattctggtgataaacaataattttaa 4802  
QY 3841 cccctccattatgtctgtatcatatccgtttgtcccccacattcattatgattcgcagcac 3900  
|||||  
Db 4803 cccctccattatgtctgtatcatatccgtttgtcccccacattcattatgattcgcagcac 4862  
QY 3901 tcttcgatgtgagttgagtggtgctgtatttaataatggtgacagatgcatctgggaatgga 3960  
|||||  
Db 4863 tcttcgatgtgagttgagtggtgctgtatttaataatggtgacagatgcatctgggaatgga 4922  
QY 3961 gagtaagaacatatacagatgcacagattactgtcttccattcattccaataatgtttgc 4020  
|||||  
Db 4923 gagtaagaacatatacagatgcacagattactgtcttccattcattccaataatgtttgc 4982  
QY 4021 cactgtgtcccttcaaaaagctgacttccctcccaaggagagatgattcctctggagacc 4080  
|||||  
Db 4983 cactgtgtcccttcaaaaagctgacttccctcccaaggagagatgattcctctggagacc 5042  
QY 4081 tcaggtgaataatccaaaagaatggtgctgcaagtgtgacttccagaagaacatgaaatgcac 4140  
|||||  
Db 5043 tcaggtgaataatccaaaagaatggtgctgcaagtgtgacttccagaagaacatgaaatgcac 5102  
QY 4141 aggaagtaactactcagggagtaaaatctctgtctaccagatgatatgtgaaggattcct 4200  
|||||  
Db 5103 aggaagtaactactcagggagtaaaatctctgtctaccagatgatatgtgaaggattcct 5162  
QY 4201 catctccagaagtcagaatgagtcagatgacttcttcttccagaatgagcaagaataaa 4260  
|||||  
Db 5163 catctccagaagtcagaatgagtcagatgacttcttcttccagaatgagcaagaataaa 5222  
QY 4261 ggttttcaagggaatcaaaactccttcacacactgtgtgaactctctagaccacgct 4320  
|||||  
Db 5223 ggttttcaagggaatcaaaactccttcacacactgtgtgaactctctagaccacgct 5282  
QY 4321 actgaactgctactccttcgaattcaacccccagagttgggtgacacagatggtccctgagat 4380  
|||||  
Db 5283 actgaactgctactccttcgaattcaacccccagagttgggtgacacagatggtccctgagat 5342  
QY 4381 ggaagttctggtgctgagggcacaggaacctctactga 4417  
|||||  
Db 5343 ggaagttctggtgctgagggcacaggaacctctactga 5379

RESULT 8  
AAD00122  
ID AAD00122 standard; DNA: 4999 BP.  
XX  
AC AAD00122;  
XX  
DT 31-JUL-2000 (first entry)  
XX  
DE Recombinant adeno associated vector construct, pVmA.1cF8deltaB.  
XX  
XX Recombinant Adeno Associated Vector; rAAV; 1cF8deltaB; promoter;  
KW human Factor VIII; hFVIII; EFlalpha; human elongation factor-1alpha;  
KW human growth hormone; hGH; ITR; inverted terminal repeat; haemophilila;  
KW gene therapy; ds.  
XX  
OS Adeno associated virus.  
XX  
PN WO200023116-A1.  
XX  
PD 27-APR-2000.  
XX  
PF 19-OCT-1999; 99WO-US24495.  
XX  
PR 20-OCT-1998; 98US-0104994.

PR 24-MAR-1999; 99US-0125974.  
PR 30-JUL-1999; 99US-0364862.  
PA (AVIG-) AVIGEN INC.  
PI Couto LB, Colosi PC;  
XX WPI; 2000-339536/29.  
XX  
PT New recombinant adenovirus-associated vector, useful for gene therapy  
PT to treat hemophilia, comprises at least a portion of Factor VIII  
PT operably linked to control sequence -  
XX  
PS Example 9; Fig 6; 92pp: English.  
XX  
CC The present DNA sequence is a recombinant adenovirus-associated vector,  
CC (rAAV) construct pVmA.1cF8deltaB. This expression vector comprises the  
CC promoter, first intron (-573 to +985) of human elongation factor-1alpha  
CC (EFlalpha) gene, human Factor VIII coding sequence (hFVIII) and a  
CC polyadenylation signal from human growth hormone (hGH). This sequence is  
CC inserted between the AAV inverted terminal repeat (ITR) regions. The  
CC hFVIII coding region comprises the heavy chain gene segment with the  
CC first 57 bp encoding the 19 amino acid signal peptide, A1 and A2 domains  
CC and 5 amino acids from the N-terminus of the B domain. The light chain  
CC segment comprises the C-terminal 85 amino acids of B domain and the A3,  
CC C1 and C2 domains. Both the heavy and light chain segments are cloned  
CC into the same plasmid separated by 42 nucleotides coding for 14 residues  
CC of the B domain, that is deleted. This plasmid is operably linked to  
CC control sequences, that directs the transcription and translation of the  
CC Factor VIII gene. The adeno-associated viral vectors are used for gene  
CC therapy to treat haemophilia. This method allows prolonged expression of  
CC therapeutic levels of Factor VIII in vivo. The rAAV are used for gene  
CC therapy, because of their broad host range, safety profile and duration  
CC of expression in the infected hosts.  
XX  
SQ Sequence 4999 BP; 1375 A; 1151 C; 1155 G; 1318 T; 0 other;

Query Match 97.8%; Score 4319.8; DB 21; Length 4999;  
Best Local Similarity 99.0%; Pred. No. 0;  
Matches 4373; Conservative 0; Mismatches 2; Indels 42; Gaps 1;

QY 1 catgcaaatgagctctccacactgctctctctgtgctcttggatctgctttagtgc 60  
|||||  
Db 415 catgcaaatgagctctccacactgctctctctgtgctcttggatctgctttagtgc 474  
QY 61 caccagaagatctactcctggtgagtggaactgtcatctggactatataatgcaagtatct 120  
|||||  
Db 475 caccagaagatctactcctggtgagtggaactgtcatctggactatataatgcaagtatct 534  
QY 121 cgtgtgagctgctgtgtgagcgaagatttctctcagaagtgccaaatctttccattcaa 180  
|||||  
Db 535 cgtgtgagctgctgtgtgagcgaagatttctctcagaagtgccaaatctttccattcaa 594  
QY 181 cactcagtcgtgtacaaaagaactggtgtgtgaattcaacggttcaacctttcaaat 240  
|||||  
Db 595 cactcagtcgtgtacaaaagaactggtgtgtgaattcaacggttcaacctttcaaat 654  
QY 241 cgtctaagcacaagggcaccctgagatggtctgtcagtcctaccatccaggtgagttta 300  
|||||  
Db 655 cgtctaagcacaagggcaccctgagatggtctgtcagtcctaccatccaggtgagttta 714  
QY 301 tgatacagtggtcattacacttaagaacatggtctccacactgtcagttctatgtct 360  
|||||  
Db 715 tgatacagtggtcattacacttaagaacatggtctccacactgtcagttctatgtct 774  
QY 361 tgggtatcctactcgtgaagaactcttgaggagctggaatgatgatcaagccaatcaag 420  
|||||  
Db 775 tgggtatcctactcgtgaagaactcttgaggagctggaatgatgatcaagccaatcaag 834  
QY 421 ggaagaagaatgataaagttctccctgtgtggaagccatacatatgtctgacagttcct 480  
|||||  
Db 835 ggaagaagaatgataaagttctccctgtgtggaagccatacatatgtctgacagttcct 894



QY	481	gaagaagatggtccaaatggtccttgacccaatgtgcctacctaatactcttcca	540
Db	895	gaagaagaaatggtccaaatggtccttgacccaatgtgccttaccatacttctcca	954
QY	541	tgtggacctgttaaagaacttgaattgaagcctatgttgagccctaccagatgtaga	600
Db	955	tgtggacctgttaaagaacttgaattgaagcctatgttgagccctaccagatgtaga	1011
QY	601	aggagatcttggccaaggaagaaagacacagaccttgcacaaatttactactttgcgt	660
Db	1015	aggagatcttggccaaggaagaaagacacagaccttgcacaaatttactactttgcgt	1074
QY	661	atttgatgaagggaagaaagtctggcactcagaacaagaacctcttgatgcagataagga	720
Db	1075	atttgatgaagggaagaaagtctggcactcagaacaagaacctcttgatgcagataagga	1134
QY	721	tgtgtgacatcgtccggccctggaccttaaatgcacagtcacatgtgtttgttaaacagtc	780
Db	1135	tgtgtgacatcgtccggccctggaccttaaatgcacagtcacatgtgtttgttaaacagtc	1194
QY	781	tcttgcagatcttgatctgatatgcacaggaatcacgtctatctgtagatgttgaaatggg	840
Db	1195	tcttgcagatcttgatctgatatgcacaggaatcacgtctatctgtagatgttgaaatggg	1254
QY	841	caaccactccatgaagtgtgaccctcaatacttcctcctgaaggtccacacattctcttggaagaa	900
Db	1255	caaccactccatgaagtgtgaccctcaatacttcctcctgaaggtccacacattctcttggaagaa	1314
QY	901	tccccaagcgctcctcttggaatctcgcacaaatcacttccctacgtcctcaaacacttgat	960
Db	1315	tccccaagcgctcctcttggaatctcgcacaaatcacttccctacgtcctcaaacacttgat	1374
QY	961	ggaaccttggaagatcttactgtttgttcacatccttccccaacaatgtaaggatgga	1020
Db	1375	ggaaccttggaagatcttactgtttgttcacatccttccccaacaatgtaaggatgga	1434
QY	1021	agcttatgttcaaaagttagacacgctgttccagaggaaaccccaactaagatgaaanaataatga	1080
Db	1435	agcttatgttcaaaagttagacacgctgttccagaggaaaccccaactaagatgaaanaataatga	1494
QY	1081	agaagaagggaagactatgatgatgtatcttactgactcttgcgaatgtgatgtgacagtttga	1140
Db	1495	agaagaagggaagactatgatgatgtatcttactgactcttgcgaatgtgatgtgacagtttga	1554
QY	1141	tgaatgaacaactcctcctccttataccaattcgtctcagttgcagaagaatctcctaaac	1200
Db	1555	tgaatgaacaactcctcctccttataccaattcgtctcagttgcagaagaatctcctaaac	1614
QY	1201	tgtggatcaacttaactatgtctctgaaggaggagatgtggagactgtcccttaagtcctgc	1260
Db	1615	tgtggatcaacttaactatgtctctgaaggaggagatgtggagactgtcccttaagtcctgc	1674
QY	1261	ccccccaatgaacagatattataaagtcacataattgaacaatgtgcccctcagcgaattgtag	1320
Db	1675	ccccccaatgaacagatattataaagtcacataattgaacaatgtgcccctcagcgaattgtag	1734
QY	1321	gaagtacaaaaaagtcgcgatttatgtgcatacacagatgaaacctttaagactgtgaagc	1380
Db	1735	gaagtacaaaaaagtcgcgatttatgtgcatacacagatgaaacctttaagactgtgaagc	1794
QY	1381	tatttaagcctgaatcagaagaaatcttggagaccttaccattatgtgggaagtgtgaagacact	1440
Db	1795	tatttaagcctgaatcagaagaaatcttggagaccttaccattatgtgggaagtgtgaagacact	1854
QY	1441	gttgaattatataatgaagaatcaagaacagacacatalaacaatctacacctcaggaattcac	1500
Db	1855	gttgaattatataatgaagaatcaagaacagacacatalaacaatctacacctcaggaattcac	1914
QY	1501	tgatgtcgtcctctgttatccaaggagattaccaaagggtgttaaacatttgaagattt	1560
Db	1915	tgatgtcgtcctctgttatccaaggagattaccaaagggtgttaaacatttgaagattt	1974

QY	1561	lccaa	ttctgcagagaaataat	ttccaataataatgaacaagtgtaagatgagcc	1620
Db	1975	lccaa	ttctgcagagaaat	taattccaataataatgaacagtgactgtagaagatgagcc	2034
QY	1621	aactaa	atcagaatccctcggtgcctgcagccgcta	ttgaactagtttcgttaatatgagag	1680
Db	2035	aactaa	atcagaatccctcggtgcctgcagccgcta	ttgaactagtttcgttaatatgagag	2094
QY	1681	agatctagcttcagagactcaat	tggccctccctcattcgtcaataagaatcgttagatca	1740	
Db	2095	agatctagcttcagagactcaat	tggccctccctcattcgtcaataagaatcgttagatca	2154	
QY	1741	aagagga	aaccaagataatgtcagaacaaagagaa	gtgcatacctggtttctcgtatttgatga	1800
Db	2155	aagagga	aaccaagataatgtcagaacaaagagaa	gtgcatacctggtttctcgtatttgatga	2214
QY	1801	gaacccaagatcgtgtactccaacagaa	tataaagaagctttccccaatccagctgagat	1860	
Db	2215	gaacccaagatcgtgtactccaacagaa	tataaagaagctttccccaatccagctgagat	2274	
QY	1861	gaagcttgaagatcccaagat	ttccaagctcccaacatcagatcaagatcaatgtagt	1920	
Db	2275	gaagcttgaagatcccaagat	ttccaagctcccaacatcagatcaagatcaatgtagt	2334	
QY	1921	ttttgataagtttcgaatgttcgaatgtt	gttgatgaagtgagacatcgtatactctaag	1980	
Db	2335	ttttgataagtttcgaatgttcgaatgtt	gttgatgaagtgagacatcgtatactctaag	2394	
QY	1981	catgttgagacaagatcgtactcccttc	gtcttccttccttcgttatataccttcaaacaa	2040	
Db	2385	catgttgagacaagatcgtactcccttc	gtcttccttccttcgttatataccttcaaacaa	2454	
QY	2041	aatggtctatgaagaacacactcaac	ccattccattccatcagaagaaactggttcattg	2100	
Db	2455	aatggtctatgaagaacacactcaac	ccattccattccatcagaagaaactggttcattg	2514	
QY	2101	gatvgaaaaa	cccaaggtctatgatctctggtgtgc	caacatcagaactcttcggaacaagag	2160
Db	2515	gatvgaaaaa	cccaaggtctatgatctctggtgtgc	caacatcagaactcttcggaacaagag	2574
QY	2161	catgaacgccttactgaaggttttctc	gtgtgtgacaagaacacttggatatacgaaga	2220	
Db	2575	catgaacgccttactgaaggttttctc	gtgtgtgacaagaacacttggatatacgaaga	2634	
QY	2221	caagtatagaata	atattccaacatactgtgcagataa	aaacaatgtgcattgaacaacaagag	2280
Db	2635	caagtatagaata	atattccaacatactgtgcagataa	aaacaatgtgcattgaacaacaagag	2694
QY	2281	ctctcccaagaatcaagacac	ccctcagacatgagcaaaagaat	taatgccaacccaccc	2340
Db	2695	ctctcccaagaatcaagacac	ccctcagacatgagcaaaagaat	taatgccaacccaccc	2712
QY	2341	agctctgaagaagcaca	taacgggaataaactgtactactcttc	agtcagatcaagaaga	2400
Db	2713	agctctgaagaagcaca	taacgggaataaactgtactactcttc	agtcagatcaagaaga	2772
QY	2401	aattgaactatga	tataccaatatactgtgaanaatga	agaaggaatttgaacattatga	2460
Db	2773	aattgaactatga	tataccaatatactgtgaanaatga	agaaggaatttgaacattatga	2832
QY	2461	tgaagatga	aaaaatcagaagcccccagactt	ccaanaagaacaagacacattattatgc	2520
Db	2833	tgaagatga	aaaaatcagaagcccccagactt	ccaanaagaacaagacacattattatgc	2892
QY	2521	tgcacgtgagagagctc	tggtgattatgtgattgtgtagct	ccccacatggtttctaagaanaaag	2580
Db	2893	tgcacgtgagagagctc	tggtgattatgtgattgtgtagct	ccccacatggtttctaagaanaaag	2952
QY	2581	ggcccaagatgagcaggt	ctccctcagttccaagaagaagtgtgtt	ccaaggaaatttaacatgag	2640
Db	2953	ggcccaagatgagcaggt	ctccctcagttccaagaagaagtgtgtt	ccaaggaaatttaacatgag	3012
QY	2641	ctctcttactcagaac	cccttataccgtgtgagaacataatga	acaatttgggaactccttgggccc	2700



|||||  
Db 3013 ctcccttaccagcccttataccgttgagaacataagacatttggactcccgggcc 3072  
Qy 2701 atatatagagcgaagtgtgaataatcatcagtgactttcagaaataagcccttcg 2760  
Db 3073 atataaagagcgaagtgtgaataatcatcagtgactttcagaaataagcccttcg 3132  
Qy 2761 tccctatctctctatctcagcttatcttctatgagaatcagaagcagaagcaga 2820  
Db 3133 tccctatctctctatctcagcttatcttctatgagaatcagaagcagaagcaga 3192  
Qy 2821 acctagaaaaaacttctcagcctaaatgaacccaacttacttctggaagtgcaca 2880  
Db 3193 acctagaaaaaacttctcagcctaaatgaacccaacttacttctggaagtgcaca 3252  
Qy 2881 tcatatggcccccaataaagatgatttgcctgaaagccttggtctatctctgagt 2940  
Db 3253 tcatatggcccccaataaagatgatttgcctgaaagccttggtctatctctgagt 3312  
Qy 2941 tgaacctgaaaaagatgtgacccagccgtgattggacccctctggtctgcacaataa 3000  
Db 3313 tgaacctgaaaaagatgtgacccagccgtgattggacccctctggtctgcacaataa 3372  
Qy 3001 cacactgaaccctgtcatgtgagacagtgacagtgacaggaatttgcctgttttcac 3060  
Db 3373 cacactgaaccctgtcatgtgagacagtgacagtgacaggaatttgcctgttttcac 3432  
Qy 3061 catcttgtatgagaccaaagctgtgatacttcaactgaaataatggaagaaactgcaggc 3120  
Db 3433 catcttgtatgagaccaaagctgtgatacttcaactgaaataatggaagaaactgcaggc 3492  
Qy 3121 tccctcgaatcagatgagaatcccaattttaaagagaattatgccttcacatgaaat 3180  
Db 3493 tccctcgaatcagatgagaatcccaattttaaagagaattatgccttcacatgaaat 3552  
Qy 3181 caatggtctacataatgatacactacactgtgctagttaatgtgcacagatcaaaagattcg 3240  
Db 3553 caatggtctacataatgatacactacactgtgctagttaatgtgcacagatcaaaagattcg 3612  
Qy 3241 atgtgtctgtctagatgtggcgcaatgaaacatccatctatctatcttcaagtgcaca 3300  
Db 3613 atgtgtctgtctagatgtggcgcaatgaaacatccatctatcttcaagtgcaca 3672  
Qy 3301 tgtgttcaactgtacgaaaaaagagagatataaatggtgcactacatctcatccagg 3360  
Db 3673 tgtgttcaactgtacgaaaaaagagagatataaatggtgcactacatctcatccagg 3732  
Qy 3361 tgtttttgagacagtgtgaaatgtttacatcccaagctgtgaatttggcggttggaatgct 3420  
Db 3733 tgtttttgagacagtgtgaaatgtttacatcccaagctgtgaatttggcggttggaatgct 3792  
Qy 3421 tatgtgcagacatctacatgctgtgagacacactttctcgtgtgtacaagcaataatg 3480  
Db 3793 tatgtgcagacatctacatgctgtgagacacactttctcgtgtgtacaagcaataatg 3852  
Qy 3481 tccagactccctctggagatggtctctcagacacatagaagatttccagattacagcttcagg 3540  
Db 3853 tccagactccctctggagatggtctctcagacacatagaagatttccagattacagcttcagg 3912  
Qy 3541 acaatatgagcagtgtggccccaaagctgtgcagactcatatctcggatcaatcaatgc 3600  
Db 3913 acaatatgagcagtgtggccccaaagctgtgcagactcatatctcggatcaatcaatgc 3972  
Qy 3601 ctgtgagcacaagaagccttctctgtgacaaagtgagatctgtgtgcaccaatgatatt 3660  
Db 3973 ctgtgagcacaagaagccttctctgtgacaaagtgagatctgtgtgcaccaatgatatt 4032  
Qy 3661 tcaagcgatcaagaaccagaggtgcccgtcagaagttctcagcctctacatctcagtt 3720  
Db 4033 tcaagcgatcaagaaccagaggtgcccgtcagaagttctcagcctctacatctcagtt 4092  
Qy 3721 tatcatcatgtatagttcttgaatggaagagtgtgcagacttatcgaagaaattccactgg 3780  
|||||

Db 4093 tatcatcatgtatagttcttgaatggaagagtgtgcagacttatcaggaattccactgg 4152  
Qy 3781 aacctaatgtctctcttggcaatgtggaattcatcggggaataaacaatatatttaa 3840  
Db 4153 aacctaatgtctctcttggcaatgtggaattcatcggggaataaacaatatatttaa 4212  
Qy 3841 cccctcaattatgtctgatacatcgtcttgacccaacatcatatagcatctcgaagcac 3900  
Db 4213 cccctcaattatgtctgatacatcgtcttgacccaacatcatatagcatctcgaagcac 4272  
Qy 3901 tcttcgcatgagatgtgagtgtgtgtatataatagttgcagaatgcattgggaatgga 3960  
Db 4273 tcttcgcatgagatgtgagtgtgtgtatataatagttgcagaatgcattgggaatgga 4332  
Qy 3961 gagtaagaacatcatcagatgcagagatctgacttcaacttccacttccatatagtttgc 4020  
Db 4333 gagtaagaacatcatcagatgcagagatctgacttcaacttccacttccatatagtttgc 4392  
Qy 4021 cactgtgtctcctcacaagaactgcacttcaactcacaaggagagatgtgctgagacc 4080  
Db 4393 cactgtgtgtcctcctcacaagaactgcacttcaactcacaaggagagatgtgctgagacc 4452  
Qy 4081 tcaagtgataatcccaaaagatgtgtcgaagtgtgacttccagaagaacaatgaaagtccac 4140  
Db 4453 tcaagtgataatcccaaaagatgtgtcgaagtgtgacttccagaagaacaatgaaagtccac 4512  
Qy 4141 aggaagttaacttaaggaatataaactctgtcttaacagcatgtatgtgaagagttcct 4200  
Db 4513 aggaagttaacttaaggaatataaactctgtcttaacagcatgtatgtgaagagttcct 4572  
Qy 4201 cactccagcagatcaagaatgcccacagtgactctcttcttcaagaatgcaagaataaa 4260  
Db 4573 cactccagcagatcaagaatgcccacagtgactctcttcttcaagaatgcaagaataaa 4632  
Qy 4261 ggttttcaaggaatacaagaactcttcaacactgtgtgtgaactcttaagccacaggt 4320  
Db 4633 ggttttcaaggaatacaagaactcttcaacactgtgtgtgaactcttaagccacaggt 4692  
Qy 4321 actgactcgttcaacttgaatttcccccagaagttgtgagacagaattgccctgaagat 4380  
Db 4693 actgactcgttcaacttgaatttcccccagaagttgtgagacagaattgccctgaagat 4752  
Qy 4381 ggaagttctgtgctgcagagcacaggaactctactga 4417  
Db 4753 ggaagttctgtgctgcagagcacaggaactctactga 4789  
  
RESULT 9  
AAA49232  
ID AAA49232 standard; DNA; 12445 BP.  
XX  
AC AAA49232;  
XX  
DT 26-SEP-2000 (first entry)  
XX  
DE Vector HSReneo for transforming endothelial cells.  
XX  
XX Hemostatic; antianemic; antidiabetic; nootropic; neuroprotective; vector;  
KW osteopathic; antislaking; immunostimulant; gene therapy; collage;  
KW endothelial cell; peripheral blood; buffy coat cell; VEGF; PCR primer;  
KW vascular endothelial growth factor; bovine brain extract; haemophilia;  
KW Factor VIII; human; transgene; adenovine deaminase deficiency; ss;  
KW sickle cell anemia; thalassemia; diabetes; alpha-antitrypsin deficiency;  
KW Alzheimer's disease; brain disease; heart disease; immune system defect;  
KW bone fracture; osteoporosis.  
XX  
OS Homo sapiens.  
OS Synthetic.  
XX  
XX WO200032750-A1.  
XX  
XX 08-JUN-2000.  
XX



PF 24-NOV-1999; 99WC-US28033.  
XX  
PR 24-NOV-1998; 98US-0109687.  
XX  
PA (MINU ) UNIV MINNESOTA.  
PA (UYEM-) UNIV EMORY.  
PA (HEBB/) HEBBEL R P.  
PA (LINY/) LIN Y.  
PA (LOLL/) LOLLAR J S.  
XX  
PI HEBBEL RP, Lin Y, Lollar JS.  
DR WPI: 2000-412303/35.  
XX  
PT Expanding population of endothelial cells useful to biocompatibilize  
PT implantable medical devices comprises contacting buffy coat cells with  
PT collagen I coated surface in culture medium comprising vascular  
PT endothelial growth factor -  
XX  
XX  
PS Claim 19; Fig 4; 53pp; English.  
XX  
XX The invention relates to a method for expanding the population of  
CC endothelial cells (EC) obtained from peripheral blood, by culturing,  
CC in contact with a collagen I coated surface, buffy coat cells obtained  
CC from peripheral mammalian blood in the presence of a culture medium  
CC containing vascular endothelial growth factor (VEGF) and free of bovine  
CC brain extract. EC are useful for treating hemophilia by introducing it  
CC into the blood stream of a mammal, so that an effective amount of  
CC Factor VIII protein is secreted in the blood stream of the mammal.  
CC This sequence represents the vector HSPneo where the human factor  
CC VIII gene into which an enhanced green fluorescent protein coding  
CC sequence (H8Q) has been inserted, is subcloned. Transgenic EC transduced  
CC in vitro are useful for improving prosthetic implants. EC is also useful  
CC for diagnosing clotting disorders where indication or disease is  
CC associated with a reduction in the activity of an enzyme. EC is also  
CC useful in gene therapy for treating the variety of diseases including  
CC adenosine deaminase deficiency, sickle cell anemia, thalassemia,  
CC hemophilia, diabetes, alpha-antitrypsin deficiency, brain disorders  
CC such as Alzheimer's disease, heart diseases, defects in immune system,  
CC for repairing bone fractures and to treat or prevent osteoporosis.  
XX  
XX  
SQ Sequence 12445 BP; 3255 A; 2989 C; 2951 G; 3250 T; 0 other:  
  
Query Match 97.7%; Score 4315; DB 21; Length 12445;  
Best Local Similarity 98.9%; Pred. No. 0;  
Matches 4370; Conservative 0; Mismatches 5; Indels 42; Gaps 1;  
  
QY 1 catgcaatagagctctccacgtctctctctgtgctcttgagattcgtcttaagtc 60  
DB 767 catgcaatagagctctccacgtctctctctgtgctcttgagattcgtcttaagtc 826  
QY 61 caccagaagatctactctgtgagtgagatgacatgacagactatagcaagagatct 120  
DB 827 caccagaagatctactctgtgagtgagatgacatgacagactatagcaagagatct 886  
QY 121 cggtagagctgctgtagcagcaaatctccctcagtagtgccaatcttttccattca 180  
DB 887 cggtagagctgctgtagcagcaaatctccctcagtagtgccaatcttttccattca 946  
QY 181 caccctagctgtagcaaaaaagactcgtttgttagaatccaggttcaacctttacaac 240  
DB 947 caccctagctgtagcaaaaaagactcgtttgttagaatccaggttcaacctttacaac 1006  
QY 241 cgttaagcagaagccaccctgagatgggtctgtagtccctcaccagagctggaagtta 300  
DB 1007 cgttaagcagaagccaccctgagatgggtctgtagtccctcaccagagctggaagtta 1066  
QY 301 tgaatacagtgtagtcaatcaatcaatcaatcaatcaatcaatcaatcaatcaatcaat 360  
DB 1067 tgaatacagtgtagtcaatcaatcaatcaatcaatcaatcaatcaatcaatcaatcaat 1126  
QY 361 tggtagtactactcagaaagctctgagagagctgcaataatgtagatcagacagtcgaag 420

DB 1127 tggtagtactactcagaaagctctgagagagctgcaataatgtagatcagacagtcgaag 1186  
QY 421 ggaagaagaagatgataaagctctccctggtggaagccatataatgctgagagctcct 480  
DB 1187 ggaagaagaagatgataaagctctccctggtggaagccatataatgctgagagctcct 1246  
QY 481 gaaagaagatggtcccaatgagcctctgacccactgtgcttactactactacttcttca 540  
DB 1247 gaaagaagatggtcccaatgagcctctgacccactgtgcttactactactacttcttca 1306  
QY 541 tgtggaacctggtataaagacttgaatccaagcctcaattggaagcctactagatagata 600  
DB 1307 tgtggaacctggtataaagacttgaatccaagcctcaattggaagcctactagatagata 1366  
QY 601 agggagcttgccaaaggaagaaagacagacacttgcacaaatttactactcttctgtc 660  
DB 1367 agggagcttgccaaaggaagaaagacagacacttgcacaaatttactactcttctgtc 1426  
QY 661 attgtagaagggaagaaagtgtgcaactcagaacaaagactccttgatgacagataagga 720  
DB 1427 attgtagaagggaagaaagtgtgcaactcagaacaaagactccttgatgacagataagga 1486  
QY 721 tgtgtcactctgctgggcttggtccttaaaatgcacacagatcaatggttataaagagtc 780  
DB 1487 tgtgtcactctgctgggcttggtccttaaaatgcacacagatcaatggttataaagagtc 1546  
QY 781 tctgccaagctcagatgtagtccacaggaagaaatgcattgagatgattggaatggg 840  
DB 1547 tctgccaagctcagatgtagtccacaggaagaaatgcattgagatgattggaatggg 1606  
QY 841 caccactctgaagtgcacccaataatctccgaaggtcacacattcttctgtgaagaaaca 900  
DB 1607 caccactctgaagtgcacccaataatctccgaaggtcacacattcttctgtgaagaaaca 1666  
QY 901 tgcgcaagcgtccttggaatctcgcgaataacttcttactgtcctaaacacttgat 960  
DB 1667 tgcgcaagcgtccttggaatctcgcgaataacttcttactgtcctaaacacttgat 1726  
QY 961 ggaacttggaagcttctcactggtttgtgcatatccttccacaaatgtagtcagga 1020  
DB 1727 ggaacttggaagcttctcactggtttgtgcatatccttccacaaatgtagtcagga 1786  
QY 1021 agcttatgtcaaatgtagcaagctgtccagaggaaccccaactacgaatgtaaaaaatga 1080  
DB 1787 agcttatgtcaaatgtagcaagctgtccagaggaaccccaactacgaatgtaaaaaatga 1846  
QY 1081 agaaagcgaagactatgtagatgtagtcttactgagatctgaaatgtagtggcaggttga 1140  
DB 1847 agaaagcgaagactatgtagatgtagtcttactgagatctgaaatgtagtggcaggttga 1906  
QY 1141 tgaatgaacactctcctctcttataccaatcgctcgaattgccaagaagcactctaaac 1200  
DB 1907 tgaatgaacactctcctctcttataccaatcgctcgaattgccaagaagcactctaaac 1966  
QY 1201 ttggttatattacattgtctgtcgaagagagagactgggactatgctccctagtcctgc 1260  
DB 1967 ttggttatattacattgtctgtcgaagagagagactgggactatgctccctagtcctgc 2026  
QY 1261 ccccgatgacgaaggttataaagatcaatattgaaacaatgagccctcagcgaattgtag 1320  
DB 2027 ccccgatgacgaaggttataaagatcaatattgaaacaatgagccctcagcgaattgtag 2086  
QY 1321 gaaatgacaanaaagtcgaattatgagcatacagatgaaaccccttaagactcgtgaag 1380  
DB 2087 gaaatgacaanaaagtcgaattatgagcatacagatgaaaccccttaagactcgtgaag 2146  
QY 1381 tattcaagcatgaatcagaatccttgggaacttacttattgagggaagltggaagacact 1440  
DB 2147 tattcaagcatgaatcagaatccttgggaacttacttattgagggaagltggaagacact 2206  
QY 1441 gttgattatatttaagaatcaagcagaagcacaataataacttacccttaacggaatcac 1500



Dh 2207 gttagattataattaagaatcaagaacagacacatatacatctacccacgaatcac 2266  
QY 1501 tggatgcgcgtccctttgtattcaagggagattaccaaaagtgtaaaaatttgaaggtt 1560  
Db 2267 tggatgcgcgtccctttgtattcaagggagattaccaaaagtgtaaaaatttgaaggtt 2226  
QY 1561 tcaaatctgcgcgggaaatattcaaatataatgtgacagtgacgttgaagaatggcc 1620  
Db 2327 tcaaatctgcgcgggaaatataatcaaatatgacagtgacgttgaagaatggcc 2386  
QY 1621 aactaaatcagatcctgcgtgtccctgacccgctla tcaactcagttctgttaataatggag 1680  
Db 2387 aactaaatcagatcctgcgtgtccctgacccgctla tcaactcagttctgttaataatggag 2446  
QY 1681 agatctcagctcagagatcatgtgcccctcctccatcctgctataaagaatcgtatgaca 1740  
Db 2447 agatctcagctcagagatcatgtgcccctcctccatcctgctataaagaatcgtatgaca 2506  
QY 1741 aagaggaaacacgataatgtccagacaagaagaaatgtcatcctgttttctgtattgtga 1800  
Db 2507 aagaggaaacacgataatgtccagacaagaagaaatgtcatcctgttttctgtattgtga 2566  
QY 1801 gaaccgaagctgtgtacctcaacagaaatatacaacgcttctccccaatccagctggagt 1860  
Db 2567 gaaccgaagctgtgtacctcaacagaaatatacaacgcttctccccaatccagctggagt 2626  
QY 1861 gcaagcttgagatccagaggttccaagcctccaacatcatgtcacagatcaaatgtgtatgt 1920  
Db 2627 gcaagcttgagatccagaggttccaagcctccaacatcatgtcacagatcaaatgtgtatgt 2686  
QY 1921 tttagtagtttcagattgttcagtttgttctgtacgtgaggtgcatactgtgacattctaa 1980  
Db 2687 tttagtagtttcagattgttcagtttgttctgtacgtgaggtgcatactgtgacattctaa 2746  
QY 1981 catgtgagcacagactgtactccttctcgtctctctctctgtgataataccttcaaacaa 2040  
Db 2747 catgtgagcacagactgtactccttctcgtctctctctgtgataataccttcaaacaa 2806  
QY 2041 aatgtgtctatgaagaacacatctcaacccctattccattctccagggaaactgttctcaatgc 2100  
Db 2807 aatgtgtctatgaagaacacatctcaacccctattccattctccagggaaactgttctcaatgc 2866  
QY 2101 gatggaaaacccaggtctcatgtgattctgtgggttgcacaactccagacttcggaaacagag 2160  
Db 2867 gatggaaaacccaggtctcatgtgattctgtgggttgcacaactccagacttcggaaacagag 2926  
QY 2161 catgaccgccttaactgaaggtttctagttgtgacaagaacactgtgtatattatcagagga 2220  
Db 2927 ctgaccgccttaactgaaggtttctagttgtgacaagaacactgtgtatattatcagagga 2986  
QY 2221 caattatgaagaatattcagcatactgtcgtgataaaaacaatgtccattgaaccaagaag 2280  
Db 2987 caattatgaagaatattcagcatactgtcgtgataaaaacaatgtccattgaaccaagaag 3046  
QY 2281 ctctcccaagaattcaagaacacccctagcactagcgaagaacaaatltatgtccaccaccc 2340  
Db 3047 ctctccctcagaat-----ccacc 3064  
QY 2341 agctctgaaaacgcacatcaacagggaaataaactcgttactactcttcagtcagatcaagaga 2400  
Db 3065 agctctgaaaacgcacatcaacagggaaataaactcgttactactcttcagtcagatcaagaga 3124  
QY 2401 aattgtactatgatataccatcatcagtttgaatgtgaagaagaagattttgacattttga 2460  
Db 3125 aattgtactatgatataccatcatcagtttgaatgtgaagaagaagattttgacattttga 3184  
QY 2461 tggagatgaanaatcagagcccccgcagcttcoaaaagaanaacacgacactattttatgtc 2520  
Db 3185 tggagatgaanaatcagagcccccgcagcttcoaaaagaanaacacgacactattttatgtc 3244  
QY 2521 tgcagttgagaaggctctggtgattatgggatagtatgtccccaatgttctlaagaaacag 2580  
Db 3245 tgcagttgagaaggctctggtgattatgggatagtatgtccccaatgttctlaagaaacag 3304

QY 2581 ggcctcagaagctggcaggtgtccctcagttccaagaagaagtggttttccaaggaaatttactatgg 2640  
Db 3305 ggcctcagaagctggcaggtgtccctcagttccaagaagaagtggttttccaaggaaatttactatgg 3364  
QY 2641 ctcccttaactcagaccctataccgttgagagaactaaatgtgaacatttggagctccggggcc 2700  
Db 3365 ctcccttaactcagaccctataccgttgagagaactaaatgtgaacatttggagctccggggcc 3424  
QY 2701 atataaagcagaagctgaagaataatcatgtgtaacttccagaatacgaagcctctcg 2760  
Db 3425 atataaagcagaagctgaagaataatcatgtgtaacttccagaatacgaagcctctcg 3484  
QY 2761 tccattatccttctatctcgtcccttatttcttgaagaagaatcagaagcgaagagaga 2820  
Db 3485 tccattatccttctatctcgtcccttatttcttgaagaagaatcagaagcgaagagaga 3544  
QY 2821 accatagaanaaactgtgtcagccctaaatgaacccaacacttaacttggaaagtgcaca 2880  
Db 3545 accatagaanaaactgtgtcagccctaaatgaacccaacacttaacttggaaagtgcaca 3604  
QY 2881 tcaatagcaccccaactaaagaatgagttgacgtcaaaagcctggcttattctctgagt 2940  
Db 3605 tcaatagcaccccaactaaagaatgagttgacgtcaaaagcctggcttattctctgagt 3664  
QY 2941 tgaagcttgaaaagaatgtgacatcagagcctgattggagcccttctgtctgcacacata 3000  
Db 3665 tgaagcttgaaaagaatgtgacatcagagcctgattggagcccttctgtctgcacacata 3724  
QY 3001 cacactgaacccctcctcatgtgagacacagtgacatcagaaatcttgcctgtttctcac 3060  
Db 3725 cacactgaacccctcctcatgtgagacacagtgacatcagaaatcttgcctgtttctcac 3784  
QY 3061 catccttgatgtgagaccaaagctgtgacttccactcgaataatgtgaagaagaactgcagggc 3120  
Db 3785 catccttgatgtgagaccaaagctgtgacttccactcgaataatgtgaagaagaactgcagggc 3844  
QY 3121 tcccttgaaatataccagaatgtgaagaatcccaacttttaagaagaatattatgcttccatgtgcaat 3180  
Db 3845 tcccttgaaatataccagaatgtgaagaatcccaacttttaagaagaatattatgcttccatgtgcaat 3904  
QY 3181 caatgtgtacataatgtgatatcactacttggcttgaatgtgtccagatcaaaagattcg 3240  
Db 3905 caatgtgtacataatgtgatatcactacttggcttgaatgtgtccagatcaaaagattcg 3964  
QY 3241 atgtgtatcgtctcagcaatgtggcagcaatgaataacatccattctatctcattcagtgagca 3300  
Db 3965 atgtgtatcgtctcagcaatgtggcagcaatgaataacatccattctatctcattcagtgagca 4024  
QY 3301 tgtgttctacgtgtacgaaaaaagaaggagttataaattgcaactgtgacatcctatcagag 3360  
Db 4025 tgtgttctacgtgtacgaaaaaagaaggagttataaattgcaactgtgacatcctatcagag 4084  
QY 3361 tgtttttgagacagtggaataatgttaccatccaagcctggaaatttggcgggtgagatgcct 3420  
Db 4085 tgtttttgagacagtggaataatgttaccatccaagcctggaaatttggcgggtgagatgcct 4144  
QY 3421 tatgtgcgagcatctacatgtcgtggatgagcaacatttctcgtgtgtacagcaataagtg 3480  
Db 4145 tatgtgcgagcatctacatgtcgtggatgagcaacatttctcgtgtgtacagcaataagtg 4204  
QY 3481 tcaagactccctctggaaatgtgcttcttgacacatataaggaattttcagaattcagaagcttccag 3540  
Db 4205 tcaagactccctctggaaatgtgcttcttgacacatataaggaattttcagaattcagaagcttccag 4264  
QY 3541 acaatatgacagtgaggccccaagaagctgtgcagagcttcatattccggatccaataatgtc 3600  
Db 4265 acaatatgacagtgaggccccaagaagctgtgcagagcttcatattccggatccaataatgtc 4324  
QY 3601 ctggagcaccaaggagccttcttctgtgataaagggtgatatcgttggcaccaatgatatt 3660  
Db 4325 ctggagcaccaaggagccttcttctgtgataaagggtgatatcgttggcaccaatgatatt 4384



QY	3661	tcgggcatcaagaccagaggtgcccgtcaagaagttctcaagcctcaacgtctctaaagt	3720
Db	4365	tcaaggcatcaagaaccagaggtgcccgtcaagaagttctcaagcctcaacgtctctaaagt	4444
QY	3721	taatacatgataagctctgatctgtaggaagaagtgagcagactlatactgaggaattccactg	3780
Db	4445	tatcatcatgataagctctgatctgtaggaagaagtgagcagactlatactgaggaattccactg	4504
QY	3781	aaccttaatgctctctcttgcgcaatgtagatcaactcgtggataaacaataattttaa	3840
Db	4505	aaccttaatgctctctctcttgcgcaatgtagatcaactcgtggataaacaataattttaa	4564
QY	3841	ccctccaattatgtgcgcgataacatccgtttgcaccccaaccattatagcattgcagac	3900
Db	4565	ccctccaattatgtgcgcgataacatccgtttgcaccccaaccattatagcattgcagac	4624
QY	3901	tcttcgcatagagttgatctggtgcctgtatataatagtctgcagcatgcattggaatgta	3960
Db	4625	tcttcgcatagagttgatctggtgcctgtatataatagtctgcagcatgcattggaatgta	4684
QY	3961	gagtaagaacatatcatcagatgacagatctactgcttaactctacttaaccaatalgttgc	4020
Db	4685	gagtaagaacatatcatcagatgacagatctactgcttaactctacttaaccaatalgttgc	4744
QY	4021	caaccgtgctctctctcaaaagcttgacttaacctcccaaggagagatgattgcttgaaacc	4080
Db	4745	caaccgtgctctctctctcaaaagcttgacttaacctcccaaggagagatgattgcttgaaacc	4804
QY	4081	tcaagttgaataatccaagaagatggtcgtgcagaatggaacttccagaagaacaatgaagaatc	4140
Db	4805	tcaagttgaataatccaagaagatggtcgtgcagaatggaacttccagaagaacaatgaagaatc	4864
QY	4141	aggagtaactactcagggaggtaaaatctctgcttaccagcatgatatgtgaaggagttcct	4200
Db	4865	aggagtaactactcagggaggtaaaatctctgcttaccagcatgatatgtgaaggagttcct	4924
QY	4201	catctccagcagatcaagaatggtccatcagtgagactctctttttccaagaatgccaagttaa	4260
Db	4925	catctccagcagatcaagaatggtccatcagtgagactctctttttccaagaatgccaagttaa	4984
QY	4261	ggtttctcagggaaatccaagaaccttccaacacctggtggaacctctagaccacagctt	4320
Db	4985	ggtttctcagggaaatccaagaaccttccaacacctggtggaacctcttagaccacagctt	5044
QY	4321	actgactcgtcaactctggaattaccaccocagagtttggtgcaaccagattgcctctgagat	4380
Db	5045	actgactcgtcaactctggaattaccaccocagagtttggtgcaaccagattgcctctgagat	5104
QY	4381	ggaggtctcgggtctgagagccaagaacctactatga 4417	
Db	5105	ggaggtctcgggtctgagagccaagaacctactatga 5141	

RESULT	10
AAD00121	
ID	AAD00121 standard; DNA; 11933 bp.
XX	
AC	AAD00121;
XX	
DT	31-JUL-2000 (first entry)
XX	
DE	Recombinant adeno associated vector construct, pAAV-F8-1.
XX	
KW	Recombinant Adeno Associated Vector; rAAV; pAAV-F8-1; human Factor VIII;
KW	hFVIII; hNF-3 albumin promoter; human elongation factor-1alpha; EF1alpha
KW	human growth hormone; hGH; Inverted terminal repeat; ITR; haemophilia;
KW	gene therapy; ds.
XX	
OS	Adeno associated virus.
XX	
PN	WO200023116-A1.
XX	
PD	27-APR-2000.

XX	19-OCT-1999;	99WO-US24495.
PF	20-OCT-1998;	98US-0104994.
PX	24-MAR-1999;	99US-0125974.
PR	30-JUL-1999;	99US-0364862.
PX	(AVIG-) AVIGEN INC.	
PA	Couto LB, Colosi PC;	
XX	WPI; 2000-339536/29.	
DR	New recombinant adenovirus-associated vector, useful for gene therapy	
PT	to treat hemophilia, comprises at least a portion of Factor VIII	
PT	operably linked to control sequence -	
XX	Example 2; Fig 5; 92pp; English.	
PS	The present DNA sequence is a recombinant adenovirus-associated vector,	
CC	(RAV) construct, pAAV-F8-1. This expression vector comprises the hNF-3	
CC	albumin promoter, the first intron (-573 to +985) of human elongation	
CC	factor-1alpha (EF1alpha) gene, human Factor VIII coding sequence (hFVIII)	
CC	and polyadenylation signal from human growth hormone (hGH). This sequence	
CC	is inserted between the AAV inverted terminal repeat (ITR) regions. The	
CC	hFVIII coding region comprises the heavy chain gene segment with the	
CC	first 57 bp encoding the 19 amino acid signal peptide, A1 and A2 domains	
CC	and 5 amino acids from the N-terminus of the B domain. The light chain	
CC	segment comprises the C-terminal 85 amino acids of B domain and the A3,	
CC	C1 and C2 domains. Both the heavy and light chain segments are cloned	
CC	into the same plasmid separated by 42 nucleotides coding for 14 residues	
CC	of the B domain, that is deleted. This plasmid is operably linked to	
CC	control sequences, that directs the transcription and translation of the	
CC	Factor VIII gene. The adeno-associated viral vectors are used for gene	
CC	therapy to treat haemophilia. This method allows prolonged expression of	
CC	therapeutic levels of factor VIII in vivo. The RAV are used for gene	
CC	therapy, because of their broad host range, safety profile and duration	
CC	of expression in the infected hosts.	
SQ	Sequence 11933 BP; 3258 A; 2818 C; 2717 G; 3140 T; 0 other;	
Query Match	96.2%; Score 4249.4; DB 21; Length 11933;	
Best Local Similarity	98.2%; Pred. No. 0;	
Matches 4338; Conservative	0; Mismatches	1; Indels 78; Gaps
OY	1 catgcaatagagcttcacactgtccttctttcttgtagcctttggagattgccttaagtc 60	
DB	426 catgcaaatagagcttcacactgtccttctttcttgtagcctttggagattgccttaagtc 485	
OY	61 caaccagaagatactactgtgtgacgtggaactgtcataatgagaaagtatcatc 120	
DB	486 caaccagaagatactactgtgtgacgtggaactgtcataatgagaaagtatcatc 545	
OY	121 cggtagagctgtcctgtggaagcgcaagattcttcctctagatgagccaaaattttccattcaa 180	
DB	546 cggtagagctgtcctgtggaagcgcaagattcttcctctagatgagccaaaattttccattcaa 605	
OY	181 gaccttcagtcggtgtacaanaaaccttgtttttagaatccagggtcaccttttcaaat 240	
DB	606 gaccttcagtcggtgtacaanaaaccttgtttttagaatccagggtcaccttttcaaat 665	
OY	241 gcctaagccaagagcccacctgtgatggtctgtttagtccctaccacagctcgagttta 300	
DB	666 gcctaagccaagagcccacctgtgatggtctgtttagtccctaccacagctcgagttta 725	
DB	726 tgatacagtggtcatcattacacttaagaacaatgcttcccaactcctgtcagcttcattgctgt 785	
OY	301 tgatacagtggtcatcattacacttaagaacaatgcttcccaactcctgtcagcttcattgctgt 360	
DB	786 tgatgtatctctactctggaagaactcttgagggagctgaatatgatgatcagacaagcttaaag 420	
DB	786 tgggttatctctactctggaagaactcttgagggagctgaatatgatgatcagacaagcttaaag 845	



QY 421 ggaagaagaatgatcaaaagtcttcctcgttggaagccatacatatgtctggcaggtcc 480  
|||||  
Db 846 gggagaaagaagatgataaagtcttcctcgttggaagccatacatatgtctggcaggtcc 905  
QY 481 gaaagaatggtgccaatggtcctctgaaccacgtgtgcttaccatcatatcttctca 540  
|||||  
Db 906 gaaagaatggtgccaatggtcctctgaaccacgtgtgcttaccatcatatcttctca 965  
QY 541 tgtggacgtgtataaagacttgaattcaggctcattggagccctactagatgtataga 600  
|||||  
Db 966 tgtggacgtgtataaagacttgaattcaggctcattggagccctactagatgtataga 1025  
QY 601 agggactctggcagaagaagacacagacctgtgcacaaattatatacttttctgt 660  
|||||  
Db 1026 agggactctggcagaagaagacacagacctgtgcacaaattatatacttttctgt 1085  
QY 661 atttgatgaagggaagaaagtgtgcactcagaaacaaagaactccttgatgaggaataga 720  
|||||  
Db 1086 atttgatgaagggaagaaagtgtgcactcagaaacaaagaactccttgatgaggaataga 1145  
QY 721 tgcctcatctgtcgggcttggtccttaaatgcaacagttcaatgttataagacagtc 780  
|||||  
Db 1146 tgcctcatctgtcgggcttggtccttaaatgcaacagttcaatgttataagacagtc 1205  
QY 781 tctggcaggtcttattgattggccacaggaatcagttcttggcagtgtattggaatgg 840  
|||||  
Db 1206 tctggcaggtcttattgattggccacaggaatcagttcttggcagtgtattggaatgg 1265  
QY 841 cacactctctgaagtgcactcaatattcctcgaaggtcacaacattcttcttgaggaca 900  
|||||  
Db 1266 cacactctctgaagtgcactcaatattcctcgaaggtcacaacattcttcttgaggaca 1325  
QY 901 tgcgacggcgtctcttggaatctcgcgaataacttccctactgtctcaaacctttgat 960  
|||||  
Db 1326 tgcgacggcgtctcttggaatctcgcgaataacttccctactgtctcaaacctttgat 1385  
QY 961 ggaaccttggaacagtttctactgttttgtaatactctccacacaatgatagtgata 1020  
|||||  
Db 1386 ggaaccttggaacagtttctactgttttgtaatactctccacacaatgatagtgata 1445  
QY 1021 agcttatgtcaaaagtagaagctgtccagaaggaacccaactaagaaatgaataatga 1080  
|||||  
Db 1446 agcttatgtcaaaagtagaagctgtccagaaggaacccaactaagaaatgaataatga 1505  
QY 1081 agaaagcggaagactatgatagtacttactgaattctgaaatgagatggtcaggttga 1140  
|||||  
Db 1506 agaaagcggaagactatgatagtacttactgaattctgaaatgagatggtcaggttga 1565  
QY 1141 tgaatgaacactcctccttcttaaccaaattgctcagttggcagaagaagatccttaaac 1200  
|||||  
Db 1566 tgaatgaacactcctccttcttaaccaaattgctcagttggcagaagaagatccttaaac 1625  
QY 1201 ttgggtacattacattgtctgtgaagaaggagctgggaactatgctccctttagtccgc 1260  
|||||  
Db 1626 ttgggtacattacattgtctgtgaagaaggagctgggaactatgctccctttagtccgc 1685  
QY 1261 ccccgatgacagaagtctataaagtcataattgaaacaatggtccctcagcgagttgag 1320  
|||||  
Db 1686 ccccgatgacagaagtctataaagtcataattgaaacaatggtccctcagcgagttgag 1745  
QY 1321 gaagtaacaaaagtcgggttttatgtgcatacacagaatgaacctttaagaactgtgaagc 1380  
|||||  
Db 1746 gaagtaacaaaagtcgggttttatgtgcatacacagaatgaacctttaagaactgtgaagc 1805  
QY 1381 tattcagcatgaatcaggaactcttggaccttacttctatggtggaggttggagacacact 1440  
|||||  
Db 1806 tattcagcatgaatcaggaactcttggaccttacttctatggtggaggttggagacacact 1865  
QY 1441 gttgatattatattaaagaatccaagaagacacataatacatctaccctcaaggaatcac 1500  
|||||  
Db 1866 gttgatattatattaaagaatccaagaagacacataatacatctaccctcaaggaatcac 1925

QY 1501 tgaatgcgtccttttattcaaggagattaccacaaaggtataaaacatttgaagattt 1560  
|||||  
Db 1926 tgaatgcgtccttttattcaaggagattaccacaaaggtataaaacatttgaagattt 1985  
QY 1561 tccaattctgcagaggaatatccaataataatgacagtgactgtagaagatggcc 1620  
|||||  
Db 1986 tccaattctgcagaggaatatccaataataatgacagtgactgtagaagatggcc 2045  
QY 1621 aactaaatcagatccctcgggtgctgagccgctattactcagtttggttataatgagag 1680  
|||||  
Db 2046 aactaaatcagatccctcgggtgctgagccgctattactcagtttggttataatgagag 2105  
QY 1681 agatctagcttcagagactcattgcccctccctcattgtgtataaagaactgtatga 1740  
|||||  
Db 2106 agatctagcttcagagactcattgcccctccctcattgtgtataaagaactgtatga 2165  
QY 1741 aagaggaacccaagataatgtcagacagaaggaatgtcaatctgttttctgtatttga 1800  
|||||  
Db 2166 aagaggaacccaagataatgtcagacagaaggaatgtcaatctgttttctgtatttga 2225  
QY 1801 gaacccaagctgtgtacctcaagagaataacaagctttctcccaatccagctggagt 1860  
|||||  
Db 2226 gaacccaagctgtgtacctcaagagaataacaagctttctcccaatccagctggagt 2285  
QY 1861 gcaagcttgaagatccagaggttccaagcctcccaatcatatgacagatcaatgtatgt 1920  
|||||  
Db 2286 gcaagcttgaagatccagaggttccaagcctcccaatcatatgacagatcaatgtatgt 2245  
QY 1921 ttcttgatgttttgcaattgttcaattgttctgcaatgaggttgcaatctgatacttcaag 1980  
|||||  
Db 2346 ttcttgatgtgttgcaattgttcaattgttctgcaatgaggttgcaatctgatacttcaag 2405  
QY 1981 cattggaagcagaactgactccttctgtctctctctctctgtgatatcaacttcaacaaa 2040  
|||||  
Db 2406 cattggaagcagaactgactccttctgtctctctctctctgtgatatcaacttcaacaaa 2465  
QY 2041 aatggtctatgaagaacacaactccatccctattctccagtgaggaactgtcttcagtgc 2100  
|||||  
Db 2466 aatggtctatgaagaacacaactccatccctattctccagtgaggaactgtcttcagtgc 2525  
QY 2101 gatggaanaacccaggtctatggaattctggtgtgccaactcagacttctggaacagag 2160  
|||||  
Db 2526 gatggaanaacccaggtctatggaattctggtgtgccaactcagacttctggaacagag 2585  
QY 2161 catgacccgcttactgaaggtttctagtgttgacaagaagaacgtgtattatacagga 2220  
|||||  
Db 2586 catgacccgcttactgaaggtttctagtgttgacaagaagaacgtgtattatacagga 2645  
QY 2221 cagttatgaagatatttcaatcactgtcgtgagtaaaacaatgccaattgaaccaagaag 2280  
|||||  
Db 2646 cagttatgaagatatttcaatcactgtcgtgagtaaaacaatgccaattgaaccaagaag 2705  
QY 2281 ctctcccggaattcaagacacccctagcacatgagcaaaagcaatttaattgcccccac 2340  
|||||  
Db 2706 ctctcccggaattcaagacacccctagcacatgagcaaaagcaatttaattgcccccac 2709  
QY 2341 agtcttgaaacgccatcaacaggaataactcgttactactcttcagttcagataagaag 2400  
|||||  
Db 2710 -----gaaataactcgttactactcttcagttcagataagaag 2747  
QY 2401 aattgactatgatataccatcagttgaatgaatgaagaaggaatttgaacttataga 2460  
|||||  
Db 2748 aattgactatgatataccatcagttgaatgaatgaagaaggaatttgaacttataga 2807  
QY 2461 tgaagatgaanaatcagagcccccagcttccaagaagaacacagacatatttata 2520  
|||||  
Db 2808 tgaagatgaanaatcagagcccccagcttccaagaagaacacagacatatttata 2867  
QY 2521 tgcagtggaagaggtcttggtgataatggaatgagtagctcccaatggttctaagaacag 2580  
|||||  
Db 2868 tgcagtggaagaggtcttggtgataatggaatgagtagctcccaatggttctaagaacag 2927  
QY 2581 ggtctcagagtggtgaggtctccctcagttccaagaaggtgtttctcaggaatttactgag 2640



Db 2928 ggcacagatggcagtgctccctcagtlccaagaagtglttccaggaattaccgagtg 2987  
Qy 2641 ctcccttaccagccctatataccgtggagaactaaatgaacattggagacccctgggccc 2700  
Db 2988 ctcccttaccagccctatataccgtggagaactaaatgaacattggagacccctgggccc 3047  
Qy 2701 atataaagagcagaagltggaagataalacatggttaacttccagaaatcagagccctcg 2760  
Db 3048 atataaagagcagaagltggaagataalacatggttaacttccagaaatcagagccctcg 3107  
Qy 2761 tccctatctcttctatcttagccttatttctatggagaataatcagagcagagcgaga 2820  
Db 3108 tccctatctcttctatcttagccttatttctatggagaataatcagagcagagcgaga 3167  
Qy 2821 acctagaanaaacttgcagcctaatagtgaaccanaaacttaacttctggaagtgcaca 2880  
Db 3168 acctagaanaaacttgcagcctaatagtgaaccanaaacttaacttctggaagtgcaca 3227  
Qy 2881 tcatatggcaccacaataagaatgaattgactgcagaagcctgggcttattctctgagt 2940  
Db 3228 tcatatggcaccacaataagaatgaattgactgcagaagcctgggcttattctctgagt 3287  
Qy 2941 tgcacttggaaaagaatgtgcactcagcctgaattggacccttctggtctgcacaccaa 3000  
Db 3288 tgcacttggaaaagaatgtgcactcagcctgaattggacccttctggtctgcacaccaa 3347  
Qy 3001 cacactgaaccctgcataatggagacaaatgacagtaacaggaattgtctgtgttttac 3060  
Db 3348 cacactgaaccctgcataatggagacaaatgacagtaacaggaattgtctgtgttttac 3407  
Qy 3061 catctttagtgaagaccaaagctgttacttcaactgnaaataatggaagaanaactgcagggc 3120  
Db 3408 catctttagtgaagaccaaagctgttacttcaactgnaaataatggaagaanaactgcagggc 3467  
Qy 3121 tccctggaatatccagatgagagatcccaactttaaagagatatgcctccatggacat 3180  
Db 3468 tccctggaatatccagatgagagatcccaactttaaagagatatgcctccatggacat 3527  
Qy 3181 caatggtacataaataatgatacaactacactcgttagtaatggtcgaagatccaagattcg 3240  
Db 3528 caatggtacataaataatgatacaactacactcgttagtaatggtcgaagatccaagattcg 3587  
Qy 3241 atgtgatctgtcgaacatggtgagcagaatgaacatccatctctatcttccagtgcaga 3300  
Db 3588 atgtgatctgtcgaacatggtgagcagaatgaacatccatctctatcttccagtgcaga 3647  
Qy 3301 tgtgttcaactgtacgaaaagaagaggtataaaatgacatgtaacatctctacaggg 3360  
Db 3648 tgtgttcaactgtacgaaaagaagaggtataaaatgacatgtaacatctctacaggg 3707  
Qy 3361 tgtttttgagacagtggaatgtttacacatccaagctggaatttggcggtgcgaatgctc 3420  
Db 3708 tgtttttgagacagtggaatgtttacacatccaagctggaatttggcggtgcgaatgctc 3767  
Qy 3421 taattgacagatctacatgctggtgagtcagacacatttctgtgttacaagaataagtg 3480  
Db 3768 taattgacagatctacatgctggtgagtcagacacatttctgtgttacaagaataagtg 3827  
Qy 3481 tccagactcccccctgggaatgtgcttctgacacatatagagatttcagatataagcttcagg 3540  
Db 3828 tccagactcccccctgggaatgtgcttctgacacatatagagatttcagatataagcttcagg 3887  
Qy 3541 acataatgacagtggtgcccacaaagctggcagacatcatatctccgatatcaatcaagc 3600  
Db 3888 acataatgacagtggtgcccacaaagctggcagacatcatatctccgatatcaatcaagc 3947  
Qy 3601 ctggagacacaaagagcccttctctggatcaaggtgtaactgtgttgacacaaatgattat 3660  
Db 3948 ctggagacacaaagagcccttctctggatcaaggtgtaactgtgttgacacaaatgattat 4007  
Qy 3661 tcaagcagataaagcccaagggtgcgcgtcagaagttctccagccctacacatctcagtt 3720

Db 4008 tcaagcagataaagcccaagggtgcgcgtcagaagttctccagccctacacatctcagtt 4067  
Qy 3721 tatcatcatgataagttcttgatggaagaagtggcagacatctatcaggaattccacttg 3780  
Db 4068 tatcatcatgataagttcttgatggaagaagtggcagacatctatcaggaattccacttg 4127  
Qy 3761 aaacctaatggtctcttcttgcaatggtgattcatctggtgataaacaacatattttaa 3840  
Db 4128 aaacctaatggtctcttcttgcaatggtgattcatctggtgataaacaacatattttaa 4187  
Qy 3841 cccctcaattatgttgatataatccgttttgcacccaactatataagcttcgcagcac 3900  
Db 4188 cccctcaattatgttgatataatccgttttgcacccaactatataagcttcgcagcac 4247  
Qy 3901 tcttcgcatggaattgataagcgttgatttaaatagttgcagcatgcatatggaaatgga 3960  
Db 4248 tcttcgcatggaattgataagcgttgatttaaatagttgcagcatgcatatggaaatgga 4307  
Qy 3961 gagtaaaagcaatalacagatgcacagatatactgcttcatcctacttaccatatgtttgc 4020  
Db 4308 gagtaaaagcaatalacagatgcacagatatactgcttcatcctacttaccatatgtttgc 4367  
Qy 4021 caccctgctctcttccaagaagctgcacttcaactccaaggagaggtatgctctggagacc 4080  
Db 4368 caccctgctctcttccaagaagctgcacttcaactccaaggagaggtatgctctggagacc 4427  
Qy 4081 tcaagtgataataatccaaagaagtgctgcagtggaattgacttccaagaagaacatgaagtcac 4140  
Db 4428 tcaagtgataataatccaaagaagtgctgcagtggaattgacttccaagaagaacatgaagtcac 4487  
Qy 4141 aggaagtaactatccaaggagtaaaatctctgcttaccagcatgataatggaaggttctc 4200  
Db 4488 aggaagtaactatccaaggagtaaaatctctgcttaccagcatgataatggaaggttctc 4547  
Qy 4201 catctccagcagtgcaagaatggtccatcgaatggagacactcttctttaaagaatggcaagtaaa 4260  
Db 4548 catctccagcagtgcaagaatggtccatcgaatggagacactcttctttaaagaatggcaagtaaa 4607  
Qy 4261 ggtttttaaaggaaatcaagaactccttccacacctgtgtggaactctctagaccacagtt 4320  
Db 4608 ggtttttaaaggaaatcaagaactccttccacacctgtgtggaactctctagaccacagtt 4667  
Qy 4321 actgaactgcgaactctggaattcaaccccagaagttgggtgcacagaattcccttgagat 4380  
Db 4668 actgaactgcgaactctggaattcaaccccagaagttgggtgcacagaattcccttgagat 4727  
Qy 4381 ggaagttctggcgtgcagggcacaggaaccttactga 4417  
Db 4728 ggaagttctggcgtgcagggcacaggaaccttactga 4764

RESULT 11  
AAN90654  
ID AAN90654 standard; DNA: 4275 BP.  
XX  
AC AAN90654:  
XX  
AC  
XX  
DT 26-JUN-1990 (first entry)  
XX  
DE DNA encoding 740 Arg-1649 Glu human Factor VIII:C.  
XX  
KW Human Factor VIII:C; Ad.RE.neo; 740 Arg-1649 Glu Factor VIII:C;  
RW haemophilila A.  
OS Homo sapiens.  
XX  
FH Key  
FT CDS 1..4275  
FT Location/Qualifiers  
FT /\*tag= a  
PN  
XX  
PD EP306968-A.  
15-MAR-1989.



XX 09-SEP-1988; 88EP-0114769.  
 PF  
 XX  
 PR 08-APR-1988; 88JP-0085454.  
 XX  
 XX  
 PA (KAGA ) CHEMO-SERO-THERAP.  
 PA (TEIJ ) TEIJUN LTD.  
 XX  
 PI Sugiyama T, Masuda K, Tajima Y, Yonemura H;  
 DR WPI; 1989-078467/11.  
 DR P-PSDB; AAP91165.  
 XX  
 PT Prodn. of recombinant human Factor-VIII-C -  
 PT using animal cells transformed with a vector contg. the gene for  
 PT Factor VIII:C and a promoter  
 XX  
 PS Fig 1(1) - 1(13); : 32pp; English.  
 XX  
 CC When translated, Arg-740 of the carboxyl terminus of the H chain is  
 CC directly bonded by a peptide bond to Glu-1649 of the amino terminus of  
 CC L chain. It is used to transform animal cells so that they produce  
 CC human Factor VIII:C. A prefd. expression vector is plasmid Ad.RE.neo.  
 CC The expression vector has at least one promoter upstream of AAP90654.  
 CC The transformants can constantly and continuously produce human Factor  
 CC VIII:C in high yield on a commercial scale. The human Factor VIII:C so  
 CC produced is considered to corresp. to the smallest species of active and  
 CC intact Factor VIII:C molecules in the human blood plasma. It is useful  
 CC for treating haemophillic A patients.  
 CC  
 XX Sequence 4275 BP; 1245 A; 941 C; 945 G; 1144 T; 0 other;



Db	1621	agagatccagcttcaggcttcatttggccctccctcccaatcctgtaacaaagacttctgagat	1680
Qy	1739	caaaaggaagaaacagatalaattgcagacaagaaggaattgcatcctgtttctgttatctgac	1798
Db	1681	caaaaggaagaaacagatalaattgcagacaagaaggaattgcatcctgtttctgttatctgac	1740
Qy	1799	gagaaacagaaagctgttacctcaacagaataatacaagctctctcccaatccagctgga	1858
Db	1741	gagaaacagaaagctgttacctcaacagaataatacaagctctctcccaatccagctgga	1800
Qy	1859	gtgcagcttggagatccagagattccaagctcccaatcatatgcagcaatcaatggcat	1918
Db	1801	gtgcagcttggagatccagagattccaagctcccaatcatatgcagcaatcaatggcat	1860
Qy	1919	gttttgaatagttttcagatttgcagatttgcgtatgcagttgagcaatactgatactta	1978
Db	1861	gttttgaatagttttcagatttgcagatttgcgtatgcagttgagcaatactgatactta	1920
Qy	1979	agcattggagacagacgactgacttcccttctgtctctctctctctgataataccttcaaac	2038
Db	1921	agcattggagacagacgactgacttcccttctgtctctctctctgataataccttcaaac	1980
Qy	2039	aaaaatgtctctatgagaacacacacacacacacacacacacacacacacacacacacac	2098
Db	1981	aaaaatgtctctatgagaacacacacacacacacacacacacacacacacacacacacac	2040
Qy	2099	tcgattgaaacacacacacacacacacacacacacacacacacacacacacacacacacac	2158
Db	2041	tcgattgaaacacacacacacacacacacacacacacacacacacacacacacacacacac	2100
Qy	2159	ggcattgacccgcttactgaaaggttctctagttgtgacaaagaacacgtgatatcaagag	2218
Db	2101	ggcattgacccgcttactgaaaggttctctagttgtgacaaagaacacgtgatatcaagag	2160
Qy	2219	gacagttatgaagatatttcagcacttctgtgagtaaaacaatgcatatgaacaaaga	2278
Db	2161	gacagttatgaagatatttcagcacttctgtgagtaaaacaatgcatatgaacaaaga	2220
Qy	2279	agcttctccagaaatccaagaacacccctagcactaggaacaaagcaatttaattgcaccca	2338
Db	2221	agcttctccagaaatccaagaacacccctagcactaggaacaaagcaatttaattgcaccca	2220
Qy	2339	ccagctcttgaaacccgcatcgaacgggaaataactcgttactcttccatccgatacaagag	2398
Db	2221	ccagctcttgaaacccgcatcgaacgggaaataactcgttactcttccatccgatacaagag	2256
Qy	2399	gaaattgactatagtatgataccataatcgaattgaaatgaaagaagaagatttgcattat	2458
Db	2257	gaaattgactatagtatgataccataatcgaattgaaatgaaagaagaagatttgcattat	2316
Qy	2459	gatgaggaatgaaataccaagagcccccagccttccaagaagaacacagacattttat	2518
Db	2317	gatgaggaatgaaataccaagagcccccagccttccaagaagaacacagacattttat	2376
Qy	2519	gtcgtcagatgaggaagagccttgagattatgaggaatgagttcccccacagttcctaagaac	2578
Db	2377	gtcgtcagatgaggaagagccttgagattatgaggaatgagttcccccacagttcctaagaac	2436
Qy	2579	agggtcagaagtgagcagttgcccctcagttcaagaagaagttgtttccagaagaatttccat	2638
Db	2437	agggtcagaagtgagcagttgcccctcagttcaagaagaagttgtttccagaagaatttccat	2496
Qy	2639	ggctccttactcagcccttataccgttggagaacataaatacaacatttggacccctgggg	2698
Db	2497	ggctccttactcagcccttataccgttggagaacataaatacaacatttggacccctgggg	2556
Qy	2699	ccatataataagagcagaagatgaaagataatacatggttaacttccagaatcagccctc	2758
Db	2557	ccatataataagagcagaagatgaaagataatacatggttaacttccagaatcagccctc	2616
Qy	2759	cgctccattcctctctattcagccttattctatgaggaagatcagaagcaagagaca	2818

---

Db	2617	ggtccctattcctctattatctagccttattctctatgaggagaatcagaagcagaagca	2676
Qy	2819	gaacctagaanaaaacttgttcaagccttaatagaacaaacttacttctggaagtgcaa	2878
Db	2677	gaacctagaanaaaacttgttcaagccttaatagaacaaacttacttctggaagtgcaa	2736
Qy	2879	catcataatggcaccacaaagaatgagttgacatgcaagaagccttggttattcttcgat	2938
Db	2737	catcataatggcaccacaaagaatgagttgacatgcaagaagccttggttattcttcgat	2796
Qy	2939	gttgacccctggaaaagaatggttgcactcagcctgagattggaacccctctgtctgcacac	2998
Db	2797	gttgacccctggaaaagaatggttgcactcagcctgagattggaacccctctgtctgcacac	2656
Qy	3059	accatcttgaatgagacccaagaagctgtgacttcaactgtaaaataatggaagaactgcag	3118
Db	2917	accatcttgaatgagacccaagaagctgtgacttcaactgtaaaataatggaagaactgcag	2976
Qy	3119	gtccctcgaatatacagatggaagatcccaactttaagaagaattatcgcttcacgca	3178
Db	2977	gtccctcgaatatacagatggaagatcccaactttaagaagaattatcgcttcacgca	3036
Qy	3179	atcaaatggtacataatgatacactacgttggttaagtatggtcagacalcaagaagat	3238
Db	3037	atcaaatggtacataatgatacactacgttggttaagtatggtcagacalcaagaagat	3096
Qy	3239	cgatggtatctgtcgaagatggtgacgaatgaanaacatccattctatcttcaatgagga	3298
Db	3097	cgatggtatctgtcgaagatggtgacgaatgaanaacatccattctatcttcaatgagga	3156
Qy	3299	catgtgttcactgttgcgaanaaaagaagatataaatggaactgttcaatctctacca	3358
Db	3157	catgtgttcactgttgcgaanaaaagaagatataaatggaactgttcaatctctacca	3216
Qy	3359	ggtgttttttgagacagtggaatgttaccatcccaagcttggaatttggcgggtggaatgc	3418
Db	3217	ggtgttttttgagacagtggaatgttaccatcccaagcttggaatttggcgggtggaatgc	3276
Qy	3419	cttatttggcgaacatctacatgctggagtgagcgaacttttctggtgttcaagaataag	3478
Db	3277	cttatttggcgaacatctacatgctggagtgagcgaacttttctggtgttcaagaataag	3336
Qy	3479	tgtaagactcccccgtggaatggtctctgagacacattagagatttccagattcaagcttca	3538
Db	3337	tgtaagactcccccgtggaatggtctctgagacacattagagatttccagattcaagcttca	3396
Qy	3539	ggaacataatggaacagtggtgcccacaaagcttgccagacttcaatttccagatcaatcaat	3598
Db	3397	ggaacataatggaacagtggtgcccacaaagcttgccagacttcaatttccagatcaatcaat	3456
Qy	3599	gocctggagcaccagaagagccttttcttggaatcgaagtgtgatacgtttggacaatgatt	3658
Db	3457	gocctggagcaccagaagagccttttcttggaatcgaagtgtgatacgtttggacaatgatt	3516
Qy	3659	attcaagcatalcaagaacccaaggttgcocgtcagaagtttcccaagccttcaatctccag	3718
Db	3517	attcaagcatalcaagaacccaaggttgcocgtcagaagtttcccaagccttcaatctccag	3576
Qy	3719	ttatcatcatatgatagttctgataatggaagaatggtgacagcttatacgaagaattccact	3778
Db	3577	ttatcatcatatgatagttctgataatggaagaatggtgacagcttatacgaagaattccact	3636
Qy	3779	ggaaccttaaatggtccttcttggcaatggtgatactcttggataaacaacataattttc	3838
Db	3637	ggaaccttaaatggtccttcttggcaatggtgatactcttggataaacaacataattttc	3696
Qy	3839	aacctccaattatgtcgcatacatcogtttggacccaactaattatacatctgcagc	3898
Db	3697	aacctccaattatgtcgcatacatcogtttggacccaactaattatacatctgcagc	3756







QY 899 catcgcaagcgctccttggaatctcgccaataacttctctactgctcaaacactcttg 958  
|||||  
Db 844 catcgcaagcgctccttggaatctcgccaataacttctctactgctcaaacactcttg 903  
QY 959 atggaacctggaagcttctactactgcttctgtaatactctctcccaacaatgctgcaatg 1018  
|||||  
Db 904 atggaacctggaagcttctactactgcttctgtaatactctctcccaacaatgctgcaatg 963  
QY 1019 gaagcttatgtccaagaagcagctgtccagaaggaaccccaactagcaatgaataaat 1078  
|||||  
Db 964 gaagcttatgtccaagaagcagctgtccagaaggaaccccaactagcaatgaataaat 1023  
QY 1079 gaagaagcgaagaactatgatactgatactctactgctgtaaatgctgctgctgctt 1138  
|||||  
Db 1024 gaagaagcgaagaactatgatactgatactctactgctgtaaatgctgctgctgctt 1083  
QY 1139 gatgatgacaactctcctctctccataccaaatcgcctcagltgccaagaagcatctctaa 1198  
|||||  
Db 1084 gatgatgacaactctcctctctccataccaaatcgcctcagltgccaagaagcatctctaa 1143  
QY 1199 acttggtatcatatcatgtgctgctgaagagagagactggtgactatgctccctctgctctc 1258  
|||||  
Db 1144 acttggtatcatatcatgtgctgctgaagagagagactggtgactatgctccctctgctctc 1203  
QY 1259 gcccccgatgacagaagtataaaagtcataatttgacaactggtccctcagcagatlgct 1318  
|||||  
Db 1204 gcccccgatgacagaagtataaaagtcataatttgacaactggtccctcagcagatlgct 1263  
QY 1319 aggaagtacaaaaagtcgcagattatggtacatacagaatgacaaccttaagaactcgtgaa 1378  
|||||  
Db 1264 aggaagtacaaaaagtcgcagattatggtacatacagaatgacaaccttaagaactcgtgaa 1323  
QY 1379 gcatctcagatgaatacagaagaatctgggaacttactttaggggaagtggggagacca 1438  
|||||  
Db 1324 gcatctcagatgaatacagaagaatctgggaacttactttaggggaagtggggagacca 1383  
QY 1439 ctgtctgatatatttaagaatccaagcagacacataataactctacccctcagcgaatc 1498  
|||||  
Db 1384 ctgtctgatatatttaagaatccaagcagacacataataactctacccctcagcgaatc 1443  
QY 1499 actgcatgctcgtctcttctgatactcaagagatctacccaaaagtgtaaaaactttgaagat 1558  
|||||  
Db 1444 actgcatgctcgtctcttctgatactcaagagatctacccaaaagtgtaaaaactttgaagat 1503  
QY 1559 ttcccaattctgcagaggaataatccaataataatggaacgtgacgtggaagatgg 1618  
|||||  
Db 1504 ttcccaattctgcagaggaataatccaataataatggaacgtgacgtggaagatgg 1563  
QY 1619 ccaactaatacagatctcctggtgcttgaccgctatactctagtttcgttaatatgag 1678  
|||||  
Db 1564 ccaactaatacagatctcctggtgcttgaccgctatactctagtttcgttaatatgag 1623  
QY 1679 agagatctagcttcaggaatcattggtccctctcctcatctgctacaagaatcgttaat 1738  
|||||  
Db 1624 agagatctagcttcaggaatcattggtccctctcctcatctgctacaagaatcgttaat 1683  
QY 1739 caaagaagaaacaggaatgctcagacaagaaggaatgctacccgttttcgttaattgat 1798  
|||||  
Db 1684 caaagaagaaacaggaatgctcagacaagaaggaatgctacccgttttcgttaattgat 1743  
QY 1799 gaagaacgaagctggtacactcctacagagaataataacaagcgtttctcccaatccagctgga 1858  
|||||  
Db 1744 gaagaacgaagctggtacactcctacagagaataataacaagcgtttctcccaatccagctgga 1803  
QY 1859 gtcagacttgagatccagagattcccaagcctcaacatcatgacagacagataatgctcat 1918  
|||||  
Db 1804 gtcagacttgagatccagagattcccaagcctcaacatcatgacagacagataatgctcat 1863  
QY 1919 gttcttgatagttgcaagttgctagttgttctgcatgaggtggcactactgttactcta 1978  
|||||  
Db 1864 gttcttgatagttgcaagttgctagttgttctgcatgaggtggcactactgttactcta 1923

QY 1979 agcatggaagcacagactgactctctcttctgctcttctctctctggaatatactctcaaac 2038  
|||||  
Db 1924 agcatggaagcacagactgactctctcttctgctcttctctctctggaatatactctcaaac 1983  
QY 2039 aaatggtctctatggaagcacactacccctattcccatctccaggaagaatgctctcag 2098  
|||||  
Db 1984 aaatggtctctatggaagcacactacccctattcccatctccaggaagaatgctctcag 2043  
QY 2099 tcgattggaanaacccaggtctctatgattctggtggtccacaactcaagactctcggaacaga 2158  
|||||  
Db 2044 tcgattggaanaacccaggtctctatgattctggtggtccacaactcaagactctcggaacaga 2103  
QY 2159 ggcattgaccgctctactgaaagttctcagttgtgacaagaacacgtgtattctcag 2218  
|||||  
Db 2104 ggcattgaccgctctactgaaagttctcagttgtgtacaagaacacgtgtattctcag 2163  
QY 2219 gacagttatggaatatttcaagcaacttctgtggtgtaaaaaatgcatctggaacaga 2278  
|||||  
Db 2164 gacagttatggaatatttcaagcaacttctgtggtgtaaaaaatgcatctggaacaga 2223  
QY 2279 agcttctccagaattcaagacacccctagacatlagcaaaaaagcaatttaagccacocca 2338  
|||||  
Db 2224 ----- 2223  
QY 2339 ccagltctgnaaacgcacatcaacggaataactcgtactactctctcagtaacatcaag 2398  
|||||  
Db 2224 ----- gaataactcgtactactctctcagtaacatcaag 2259  
QY 2399 gaattgactatgattgatacattacagttgaaatggaagaaggaatttgacattat 2458  
|||||  
Db 2260 gaattgactatgattgatacattacagttgaaatggaagaaggaatttgacattat 2319  
QY 2459 gatgagatgnaaaatcagagagcccgagactcttcaaaaagaaacacagacactattat 2518  
|||||  
Db 2320 gatgagatgnaaaatcagagagcccgagactcttcaaaaagaaacacagacactattat 2379  
QY 2519 gctcgaagtggagaggtctctggtattatggaatggtgtgtccccaactgttccaagaac 2578  
|||||  
Db 2380 gctcgaagtggagaggtctctggtattatggaatggtgtgtgtccccaactgttccaagaac 2439  
QY 2579 agggctcagagttgagaggtcctccagacttcaagaagttgtttccaggaattactgat 2638  
|||||  
Db 2440 agggctcagagttgagaggtcctccagacttcaagaagttgtttccaggaattactgat 2499  
QY 2639 ggtcctcttactcagcccttaaacgttggagaactaaatgaaacttgggaactcctggg 2698  
|||||  
Db 2500 ggtcctcttactcagcccttaaacgttggagaactaaatgaaacttgggaactcctggg 2559  
QY 2699 ccataataagaagcagaagttggaagataataatcatgtgttaacttcaagaatcaggtctct 2758  
|||||  
Db 2560 ccataataagaagcagaagttggaagataataatcatgtgttaacttcaagaatcaggtctct 2619  
QY 2759 gctccctatctcctctatctcagccttatcttcttaataaggaagaatcaggaagcaga 2818  
|||||  
Db 2620 gctccctatctcctctatctcagccttatcttcttaataaggaagaatcaggaagcaga 2679  
QY 2819 gaacctagaanaaaacttgttcaaaccttaatgaaacccaacttacttcttgaagtgcaa 2878  
|||||  
Db 2680 gaacctagaanaaaacttgttcaaaccttaatgaaacccaacttacttcttgaagtgcaa 2739  
QY 2879 catcatatggaacccaactaaagatgagttggaactgcaagacgtggtgtattctctcgtat 2938  
|||||  
Db 2740 catcatatggaacccaactaaagatgagttggaactgcaagacgtggtgtattctctcgtat 2799  
QY 2939 gttgaacctggaanaagatggtcactcagccttgaatgtgaacccctctggtctgccaact 2998  
|||||  
Db 2800 gttgaacctggaanaagatggtcactcagccttgaatgtgaacccctctggtctgccaact 2859  
QY 2999 aacacactgaaacccctgctcagcaggaacaaatggaacgtacaggaattgtcgtgttttc 3058  
|||||  
Db 2860 aacacactgaaacccctgctcagcaggaacaaatggaacgtacaggaattgtcgtgttttc 2919  
QY 3059 aacatcttctgataagacaaaagcgtgtacttcaactgaaaaataatggaagaagaactgcag 3118



```
|||||
Db 2920 accatcttgatgagaccaaaagctggtacttcaactcgtgaaataatggaagaactcgagg 2979
QY 3119 gctccctgtcaatatccagatggaagatcccaactttaagaagaatatcgctccatgca 3178
Db 2980 gctccctgtcaatatccagatggaagatcccaactttaagaagaatatcgctccatgca 3039
QY 3179 atcaatggcacaataatgtatgatacctggttagtaatgtgctgaagatccaagaatt 3238
Db 3040 atcaatggcacaataatgtatgatacctggttagtaatgtgctgaagatccaagaatt 3099
QY 3239 cgaatgtatctgctcagcagctgaggcagaatgaaacatccatctatctcaattcagtgga 3298
Db 3100 cgaatgtatctgctcagcagctgaggcagaatgaaacatccatctatctcaattcagtgga 3159
QY 3299 catgtgttcaactgtacgaaaaaagagagatataaaatggcactgtacaactctatcca 3358
Db 3160 catgtgttcaactgtacgaaaaaagagagatataaaatggcactgtacaactctatcca 3219
QY 3359 ggtgtttttgagacagtgaagaatgtaccatccaaaagcttgaatttggcggtggaatgc 3418
Db 3220 ggtgtttttgagacagtgaagaatgtaccatccaaaagcttgaatttggcggtggaatgc 3279
QY 3419 ctatttggcgagcatctacatgcttggatgagacacacttttctgtgtgtacaagaataag 3478
Db 3280 ctatttggcgagcatctacatgcttggatgagacacacttttctgtgtgtacaagaataag 3339
QY 3479 tgtcagaactccctgggaatgcttcttgacacatagagatttcaagattacagcttca 3538
Db 3340 tgtcagaactccctgggaatgcttcttgacacatagagatttcaagattacagcttca 3399
QY 3539 ggaacaataggaagtgaggccccaagctggcagaactcatctatcccgatcaatcaat 3598
Db 3400 ggaacaataggaagtgaggccccaagctggcagaactcatctatcccgatcaatcaat 3459
QY 3599 gctctgagacccaagagcccttcttcttgatcaagtgatctgttggcacaatgatt 3658
Db 3460 gctctgagacccaagagcccttcttcttgatcaagtgatctgttggcacaatgatt 3519
QY 3659 attcagcgcatcaagaaccgaaggtgcccgatcagaagttctccagccttcaatctctcag 3718
Db 3520 attcagcgcatcaagaaccgaaggtgcccgatcagaagttctccagccttcaatctctcag 3579
QY 3719 ttatatacatgatatgcttcttgggaagaatgagcagaactatcgagaattccact 3778
Db 3580 ttatatacatgatatgcttcttgggaagaatgagcagaactatcgagaattccact 3639
QY 3779 ggaaccttaatgcttcttcttgcaatgttgatctactcttggaataaacaacaattttt 3838
Db 3640 ggaaccttaatgcttcttcttgcaatgttgatctactcttggaataaacaacaattttt 3699
QY 3839 aaacctcaaatatgtctcgatacatccgtttgacccaacatcatatagatctcgagc 3898
Db 3700 aaacctcaaatatgtctcgatacatccgtttgacccaacatcatatagatctcgagc 3759
QY 3899 acctctcgatagagtttgggttggcttgaattaaatgattggcagctgcatgggaatg 3958
Db 3760 acctctcgatagagtttgggttggcttgaattaaatgattggcagctgcatgggaatg 3819
QY 3959 gagaagtaagaacaatacgaatgacagatctactgtcttcaacttcaacaataatgatt 4018
Db 3820 gagaagtaagaacaatacgaatgacagatctactgtcttcaacttcaacaataatgatt 3879
QY 4019 gccacctgtctctctcaaaagctcgactacccccaagggagagtaatgctgtgaga 4078
Db 3880 gccacctgtctctctcaaaagctcgactacccccaagggagagtaatgctgtgaga 3939
QY 4079 cctcaggttaataatacaaaagtgctgcaagtgaactccaagaagaacaataagatc 4138
Db 3940 cctcaggttaataatacaaaagtgctgcaagtgaactccaagaagaacaataagatc 3999
QY 4139 acagaggtactactcagaggaataaatctctgttaccagcatgtatgtgaaggatc 4198
|||||
```

```
Db 4000 acagagtaactactcagggagtaaaatctctgtcttaccagcaatgtatgtgaaggatc 4059
QY 4199 ctatcttccagcagctcaagtgccatcatgtgactctcttttcaagaatggaagta 4258
Db 4060 ctatcttccagcagctcaagtgccatcatgtgactctcttttcaagaatggaagta 4119
QY 4259 aaggttttccaggaatacagaactctctacaccttggtagactctcagaaccccg 4318
Db 4120 aaggttttccaggaatacagaactctctacaccttggtagactctcagaaccccg 4179
QY 4319 ttactgactcgtctactctcgaaatcccccagaagtgggtgacacagatggccctgag 4378
Db 4180 ttactgactcgtctactctcgaaatcccccagaagtgggtgacacagatggccctgag 4239
QY 4379 atggaagttctggcctggcagagcaggaactctac 4414
Db 4240 atggaagttctggcctggcagagcaggaactctac 4275
```

## RESULT 13

AAN80447  
ID AAN80447 standard; DNA; 4272 BP.XX AAN80447;  
ACXX 10-OCT-1990 (first entry)  
DTXX DE Modified factor VIII:C sequence with the R740-D1658 deletion.  
XXXX KW Modified factor VIII:C; haemophilia; procoagulant;  
XX blood coagulation; RD deletion; ss.XX OS Homo sapiens.  
OSXX PN W08800831-A.  
PNXX PD 11-FEB-1988.  
PDXX 31-JUL-1987; 87WO-US01814.  
XXXX PR 01-AUG-1986; 86US-0893375.  
XXXX PA (BIOD ) BIOGEN NV.  
PAXX PI Pasek MP;  
PIXX DR WPI; 1988-049866/07.  
DRXX P-PSDB; AAP80268.  
XXXX PT New DNA sequences encoding modified factor VIII:C - with deletion of DNA  
XX encoding maturation polypeptide, useful for high yield transformation.XX PS Claim 3; Page 47-48-49-50; 97pp; English.  
PSXX CC The RD deletion removes the DNA from Ser 741 to Ser 1657. A major part  
XX of the sequence encoding the maturation polypeptide ofXX CC factor VIII:C is deleted, i.e. Gln 744 - Asp 1563.  
XXXX CC The full length factor VIII:C cDNA has two changes with respect to the  
XX published sequence (EPO application 160457):XX CC CTG to CTA at Leu 242 and TTC to CTC change at amino acid residue 1880  
XX (Phe to Leu). The product is produced in approx. 20 times higherXX CC yields than previous recombinant produced factor VIII:C and are more  
XX easily purified. The peptide is used for treating haemophilia A, bothXX CC acute and prolonged bleeding.  
XX See also AAN80444 and AAN80446.

XX SO Sequence 4272 BP; 1243 A; 941 C; 944 G; 1144 T; 0 other;

Query Match 94.4%; Score 4170; DB 9; Length 4272;  
Best Local Similarity 98.0%; Pred. No. 0;  
Matches 4267; Conservative 0; Mismatches 5; Indels 84; Gaps 1;



OY	59	gaaaccaaagaatgactaaccggtggtgacgttggaactgcatggtgagctatataccaagtat	118
Db	1	ggccacagaagatactactaccctgggtgagctggaactgtcatactggagctatacgcaagtat	60
OY	119	ctcgctgagctgagctgtgagcagcaagaattcctctctagagtgaccaaaattcttccatc	178
Db	61	ctcgggtgagctgctgtgtgagcgcgaagaattcctctctagagtgccaaaaattcttccatc	120
OY	179	aaacactcagtcgtgatacaaaaaagactcgtttgtagaatcaccgttcacatttccac	238
Db	121	aaacactcagtcgtgatacaaaaaagactcgtttgtagaatcaccgttcacatttccac	180
OY	239	atcgcctaaagccaaagcgcaacctgtgagtgtctgtatgctcctaccacccagctgaggt	298
Db	181	atcgcctaaagccaaagcgcaacctgtgagtgtctgtatgctcctaccacccagctgaggt	240
OY	299	tatgatacagtgatgcatatacctataaagaatactgctccctccctgcaagcttcacgt	358
Db	241	tatgatacagtgatgcatatacctataaagaatactgctccctccctgcaagcttcacgt	300
OY	359	gttgtgtatctcactctgnaaagctcttgaggtgagctgataatgatactgacagccagtca	418
Db	301	gttgtgtatctcactctgnaaagctcttgaggtgagctgataatgatactgacagccagtca	360
OY	419	agggagaaagaagaatgataaagtctctccctgggtgagagccatacatatgtctgcaagtc	478
Db	361	agggagaaagaagaatgataaagtctctccctgggtgagagccatacatatgtctgcaagtc	420
OY	479	cgtgaagaagaatgtcccaatgtgctctgcagccaaatgctgcttaactcatactctct	538
Db	421	cgtgaagaagaatgtcccaatgtgctctgcagccaaatgctgcttaactcatactctct	480
OY	539	catgtgacacctgtgtaaaagacttgaaatctgagctcataltgagagcctactagtata	598
Db	481	catgtgacacctgtgtaaaagacttgaaatctgagctcataltgagagcctactagtata	540
OY	599	gaagagagatctggtccaaaggaagaagacacacacctgtgcacaaattatactactttgtc	658
Db	541	gaagagagatctggtccaaaggaagaagacacacacctgtgcacaaattatactactttgtc	600
OY	659	gatatgtgataagggaaagaatctgacatctgcaagaacaaagaactcctgtatgcagatag	718
Db	601	gatatgtgataagggaaagaatctgacatctgcaagaacaaagaactcctgtatgcagatag	660
OY	719	gatlgtcgtacatctgtccgggcctggtcgtctaaaaatgcacacagtcaaatggtatgataacag	778
Db	661	gatlgtcgtacatctgtccgggcctggtcgtctaaaaatgcacacagtcaaatggtatgataacag	720
OY	779	tctctgcagagttctgtattgtgataatgcacacagaagaatcagatcatgtgcatgtgattggaat	838
Db	721	tctctgcagagttctgtattgtgataatgcacacagaagaatcagatcatgtgcatgtgattggaat	780
OY	839	ggcacaacacccctgaagaatgacataataattcctcgaaggtcacacattctctgtagaagac	898
Db	781	ggcacaacacccctgaagaatgacataataattcctcgaaggtcacacattctctgtagaagac	840
OY	899	catcgccaggtcgtccttggaaatctcgcgaataataattcctctactgtctcaacaactctg	958
Db	841	catcgccaggtcgtccttggaaatctcgcgaataataattcctctactgtctcaacaactctg	900
OY	959	atggagacctggagagattctcaccggtttgtgcatatactctctcccaacaatgatgagcat	1018
Db	901	atggagacctggagagattctcaccggtttgtgcatatactctctcccaacaatgatgagcat	960
OY	1019	gaagactatgtcaaaatgacagactgcccagaggaaccccaactgacgaatgaaataaat	1078
Db	961	gaagactatgtcaaaatgacagactgcccagaggaaccccaactgacgaatgaaataaat	1020
OY	1079	gaagaagcggaaagactatactgatactcattacagatctcgaatctgatactgatactgatact	1138
Db	1021	gaagaagcggaaagactatactgatactcattacagatctcgaatctgatactgatactgatact	1080
OY	1139	gatgatgacaactctcctctcttataccaaattcgtctcagttgccaagaagcatcctaa	1198

Dh	1081	gatyatgcaacatccctccctccatcccaatccgcagcttgcagaagaacatcccaa	1140
Qy	1199	acttgggtacatctacatctgtctgcgaagagagacttgggtacatctctccctta	1258
Dh	1141	acttgggtacatctacatctgtctgcgaagagagacttgggtacatctctccctta	1200
Qy	1259	gcccccgtatgcagaagttataaagtcacatcttgaacaatgcccctcagcgaatg	1318
Dh	1201	gcccccgtatgcagaagttataaagtcacatcttgaacaatgcccctcagcgaatg	1260
Qy	1319	aggaagatcaaaaaagccgatattatggcaacaagatgaaacctttagagctgttaa	1378
Dh	1261	aggaagatcaaaaaagccgatattatggcaacaagatgaaacctttagagctgttaa	1320
Qy	1379	gcatctacagcatgaatcaaggaatctcttgcgaaccttctatctatggtggaagctt	1438
Dh	1321	gcatctacagcatgaatcaaggaatctcttgcgaaccttctatctatggtggaagctt	1380
Qy	1439	ctgttgtatataatttaagaatcaagaacagacacataacatctacccctcaggaat	1498
Dh	1381	ctgttgtatataatttaagaatcaagaacagacacataacatctacccctcaggaat	1440
Qy	1499	actatgtgcgcctctctgttatttcaagaggttaccaaaagtgtaaaacatttgaagat	1558
Dh	1441	actatgtgcgcctctctgttatttcaagaggttaccaaaagtgtaaaacatttgaagat	1500
Qy	1559	tttccaattctgcagaggaagaatattcaaatataatgtgacagctgtacttgaagaatg	1618
Dh	1501	tttccaattctgcagaggaagaatattcaaatataatgtgacagctgtacttgaagaatg	1560
Qy	1619	ccaactaatatcagatctctgcgtgcctgcagccgctatctactagtttcgttaatatgag	1678
Dh	1561	ccaactaatatcagatctctgcgtgcctgcagccgctatctactagtttcgttaatatgag	1620
Qy	1679	aggaatatagcttcaaggaacatctatgtgcccctctctactatctgttcaaaaagatcttgaat	1738
Dh	1621	aggaatatagcttcaaggaacatctatgtgcccctctctactatctgttcaaaaagatcttgaat	1680
Qy	1739	caaaagagaaacccagaaataatgtctacagcaagaagaaatgtcatcctgtttctgtatcttat	1798
Dh	1681	caaaagagaaacccagaaataatgtctacagcaagaagaaatgtcatcctgtttctgtatcttat	1740
Qy	1799	gagaaacccaagctgtgtacccctcacagagaatatacaacgcttctccccaatccagcttga	1858
Dh	1741	gagaaacccaagctgtgtacccctcacagagaatatacaacgcttctccccaatccagcttga	1800
Qy	1859	gtgcagagcttgaaggttccaaggttccaaagctcccaacataatgcaagaactaaatgtgcat	1918
Dh	1801	gtgcagagcttgaaggttccaaggttccaaagctcccaacataatgcaagaactaaatgtgcat	1860
Qy	1919	gtttttatagtttgcaggttgcacagtttctgttcagagagtggtgcatactgcgtgaatctcta	1978
Dh	1861	gtttttatagtttgcaggttgcacagtttctgttcagagagtggtgcatactgcgtgaatctcta	1920
Qy	1979	agcatctggaagcagagactgtacatctcctctctgtccctctctctgtgataactctcaaacac	2038
Dh	1921	agcatctggaagcagagactgtacatctcctctctgtccctctctctgtgataactctcaaacac	1980
Qy	2039	aaatattgtctatgaaagcacaactcaccctatctccattctcagggagaactgttctctatg	2098
Dh	1981	aaatattgtctatgaaagcacaactcaccctatctccattctcagggagaactgttctctatg	2040
Qy	2099	tcgaatgaaacccaaggtctcatgattctcgtgggtgcacaacatccagaacttctggagaacga	2158
Dh	2041	tcgaatgaaacccaaggtctcatgattctcgtgggtgcacaacatccagaacttctggagaacga	2100
Qy	2159	ggcaatgacgcgccttactgaaaggttctctagtctgtgacaagaacacgtgtgatattacagag	2218
Dh	2101	ggcaatgacgcgccttactgaaaggttctctagtctgtgacaagaacacgtgtgatattacagag	2160
Qy	2219	gacacgttatgaaagatattccaagcatctgtgcgtgagtaaaaacaatgtccaattgaaaccaaga	2278



Dh 2161 gacagatataatattcaagatacttgctagtaaaaaaatgcatatgaaccaaga 2220  
Qy 2279 agcttcccaagattcaagaacacctagcaactagcaaaagaattatgcccacca 2338  
Db 2221 ----- 2220  
Qy 2339 ccagcttgaaacgcatcaacaggaataaactcgtactacttctcagtcagaaga 2398  
Db 2221 -----gaataaactcgtactacttctcagtcagaaga 2256  
Qy 2399 gaattgactatgataaccatacaagttagaatagaagaagaatttgacattat 2458  
Db 2257 gaattgactatgataaccatacaagttagaatagaagaagaatttgacattat 2316  
Qy 2459 gatgagatgaataatcagaagccccgcagcttcaagaagaacagacactttatt 2518  
Db 2317 gatgagatgaataatcagaagccccgcagcttcaagaagaacagacactttatt 2376  
Qy 2519 gctgcagtgagagagctctgagatlaaggatagtagtcccaacatgltctaaagaac 2578  
Db 2377 gctgcagtgagagagctctgagatlaaggatagtagtcccaacatgltctaaagaac 2436  
Qy 2579 aaggctcagagtgagagtgctccctcagttcaagaagaagtgtttccagaagaattactgat 2638  
Db 2437 aaggctcagagtgagagtgctccctcagttcaagaagaagtgtttccagaagaattactgat 2496  
Qy 2639 ggcctcttactagcccttatccgttgagaactaaatgaagaatttgagactcccgagg 2698  
Db 2497 ggcctcttactagcccttatccgttgagaactaaatgaagaatttgagactcccgagg 2556  
Qy 2699 ccataataagacaagaattgagaataatatacatagtaacttccagaatacagagctct 2758  
Db 2557 ccataataagacaagaattgagaataatatacatagtaacttccagaatacagagctct 2616  
Qy 2759 cgtccctatctctatctctagcttacttctatgagaagaatcagaagcagaagca 2818  
Db 2617 cgtccctatctctatctctagcttacttctatgagaagaatcagaagcagaagca 2676  
Qy 2819 gaacctagaagaacttctcaagcctaatagaacaaacttcttggaagaagcgca 2878  
Db 2677 gaacctagaagaacttctcaagcctaatagaacaaacttcttggaagaagcgca 2736  
Qy 2879 catcataatgcaaccactaaagaatgagtttgactgcaaaagcctgagcttatctctgat 2938  
Db 2737 catcataatgcaaccactaaagaatgagtttgactgcaaaagcctgagcttatctctgat 2796  
Qy 2939 gttgacctggaagaagaatgctcaatcagagcctgattggaacccctctggtctgcaact 2998  
Db 2797 gttgacctggaagaagaatgctcaatcagagcctgattggaacccctctggtctgcaact 2856  
Qy 2999 aacacactgaaccctggtcatgtgagacaagtgacagtagtaagaatttgcctggttttc 3058  
Db 2857 aacacactgaaccctggtcatgtgagacaagtgacagtagtaagaatttgcctggttttc 2916  
Qy 3059 accatcttgatgagaccaaagcgtgtacttcaactgaataataatggaagaacatgcag 3118  
Db 2917 accatcttgatgagaccaaagcgtgtacttcaactgaataataatggaagaacatgcag 2976  
Qy 3119 gctccctgcaataatccagaatggaagaatcccaactttaaagaagaatatacgtctcactga 3178  
Db 2977 gctccctgcaataatccagaatggaagaatcccaactttaaagaagaatatacgtctcactga 3036  
Qy 3179 atcaaatggtataataatgatacactactggttagtaatggtcagaatcaaaagatt 3238  
Db 3037 atcaaatggtataataatgatacactactggttagtaatggtcagaatcaaaagatt 3096  
Qy 3239 cgatggtatctgctcagcagctgggcaagcaatgaaacatcatctctatctcaatgga 3298  
Db 3097 cgatggtatctgctcagcagctgggcaagcaatgaaacatcatctctatctcaatgga 3156  
Qy 3299 catggttcaactgctacgaaaaaagagagataaaatggcagctgtacaactctcactca 3358  
Db 3157 catggttcaactgctacgaaaaaagagagataaaatggcagctgtacaactctcactca 3216

Qy 3359 ggtgttttgagacagtggaatggttaccatcccaagcttggaatttggcggtggaatgc 3418  
Db 3217 ggtgttttgagacagtggaatggttaccatcccaagcttggaatttggcggtggaatgc 3276  
Qy 3419 ctatctgagagcatctacatgcttgagatgagacacttcttctggtgtgacagaataag 3478  
Db 3277 ctatctgagagcatctacatgcttgagatgagacacttcttctggtgtgacagaataag 3336  
Qy 3479 tgtcagaactcccttggaatggttcttgacaacatgaagaatttcaagatlaagcttca 3538  
Db 3337 tgtcagaactcccttggaatggttcttgacaacatgaagaatttcaagatlaagcttca 3396  
Qy 3539 ggaacaatgagacagtgggccccaagctggcgagacttcaattccagatcaatcaat 3598  
Db 3397 ggaacaatgagacagtgggccccaagctggcgagacttcaattccagatcaatcaat 3456  
Qy 3599 gcttgagacccaagagcccttctcttgatcaaggttgatctgttggcaccatgat 3658  
Db 3457 gcttgagacccaagagcccttctcttgatcaaggttgatctgttggcaccatgat 3516  
Qy 3659 attcagcatcaagaaccaaggtgcccgtcagaagttctcagccttcaatctcag 3718  
Db 3517 attcagcatcaagaaccaaggtgcccgtcagaagttctcagccttcaatctcag 3576  
Qy 3719 ttatcatatgatatgcttcttgatggagaagtggcagactatcagagaatttccact 3778  
Db 3577 ttatcatatgatatgcttcttgatggagaagtggcagactatcagagaatttccact 3636  
Qy 3779 ggaaccttaatgcttctcttgacaatgctgcatctcgggataaacaacatattt 3838  
Db 3637 ggaaccttaatgcttctcttgacaatgctgcatctcgggataaacaacatattt 3696  
Qy 3839 aacctccaattatgctcgtatcatatccgtttgacccaactcaattatagcatctgcag 3898  
Db 3697 aacctccaattatgctcgtatcatatccgtttgacccaactcaattatagcatctgcag 3756  
Qy 3899 actctcgtatgagtgatgagctggcgttgatggaattgaattgagcagctgcagtggaatg 3958  
Db 3757 actctcgtatgagtgatgagctggcgttgatggaattgaattgagcagctgcagtggaatg 3816  
Qy 3959 gagagtaagaacatatacagaatgacagatctactctcaacttcaacatattgtt 4018  
Db 3817 gagagtaagaacatatacagaatgacagatctactctcaacttcaacatattgtt 3876  
Qy 4019 gccacctggtctcctccaagaactcgaactcaactccaagggagagatattgcttgaga 4078  
Db 3877 gccacctggtctcctccaagaactcgaactcaactccaagggagagatattgcttgaga 3936  
Qy 4079 cctcaggtgaataatccaagaagtggtcgaagtgagacttccaagaagaatgaagaatgc 4138  
Db 3937 cctcaggtgaataatccaagaagtggtcgaagtgagacttccaagaagaatgaagaatgc 3996  
Qy 4139 acagaggttaactactcagaggttaaatctcgtcttaccagcatgatatgtgaaggttc 4198  
Db 3997 acagaggttaactactcagaggttaaatctcgtcttaccagcatgatatgtgaaggttc 4056  
Qy 4199 ctcaatctcagcagatcaagaatggtcgaatcagtggaactctcttcttcaagaatggaagta 4258  
Db 4057 ctcaatctcagcagatcaagaatggtcgaatcagtggaactctcttcttcaagaatggaagta 4116  
Qy 4259 aaggttttccaaggaataaagaactctcacaacgttggaactcttaagaaccaagc 4318  
Db 4117 aaggttttccaaggaataaagaactctcacaacgttggaactcttaagaaccaagc 4176  
Qy 4319 ttactgactcgtctacacttgaattcaaccccaaggttgggtgcacacagattccctcagag 4378  
Db 4177 ttactgactcgtctacacttgaattcaaccccaaggttgggtgcacacagattccctcagag 4236  
Qy 4379 atggaaggttcctgggtcgtgagagcagaagactctac 4414  
Db 4237 atggaaggttcctgggtcgtgagagcagaagactctac 4272



RESULT 14

AAx82258  
ID AAX82258 standard; cDNA; 4373 BP.

XX AAX82258;

DT 18-AUG-1999 (first entry)

DE Beta domain deleted Factor VIII protein encoding cDNA.

XX Factor VIII protein; gene modification; gene therapy; clinical disorder;

KW splicing pattern; RNA processing; gene regulation; beta-domain; human;

KW ss.

XX Homo sapiens.

OS WO929848-A1.

PN 17-JUN-1999.

PD 25-NOV-1998; 98MO-US5354.

PR 16-JAN-1998; 98US-0071596.

PR 05-DEC-1997; 97US-0067614.

PA (IMMU-) IMMUNE RESPONSE CORP.

PI Bidlingmaier S, Gonzales JEN, Ill CR, Yang CQ;

DR WPI; 1999-385602/32.

PT P-PSDB; AAY21675.

PT Genes and vectors exhibiting increased expression and novel splicing

PT patterns, useful for expression of, e.g. beta-domain deleted factor

PT VIII

PS Claim 9; Page 72-78; 123pp; English.

CC The invention describes novel genes and vectors exhibiting increased

CC expression and novel splicing patterns. It provides a gene encoding a

CC Factor VIII protein, that comprises one or more consensus or near

CC consensus splice sites which have been corrected to increase expression.

CC The method, DNA sequences and expression vectors can be used to increase

CC the expression of a gene, especially a Factor VIII gene. Genes containing

CC modified 5' and/or 3' untranslated regions have optimized expression

CC levels and tissue-specific expression. The methods are used for

CC identification and correction of consensus splice sites, addition of

CC introns, optimization of 5' and 3' untranslated regions and increase in

CC cytoplasmic RNA accumulation. Hence the DNAs are useful in gene therapy

CC to treat a clinical disorder, to study RNA processing and/or gene

CC regulation. The present sequence represents a cDNA encoding a beta-domain

CC deleted factor VIII protein.

XX

XX

Sequence 4373 BP; 1278 A; 1011 C; 938 G; 1146 T; 0 other:

Query Match 94.2%; Score 4161; DB 20; Length 4373;

Best Local Similarity 96.8%; Pred. No. 0; Mismatches 100; Indels 42; Gaps 1;

Matches 42/3; Conservative 0; Mismatches 100; Indels 42; Gaps 1;

QY 2 atgcaataagctctccacgtctcttctgtcgtcttgcgtatctgtcttagtgc 61

DB 1 atggaataagctctccacgtctcttctgtcgtcttgcgtatctgtcttagtgc 60

QY 62 accaagaatactactcctgggtgcagtggaactgtcatggaactatagcaagtgatctc 121

DB 61 accaagaatactactcctgggtgcagtggaactgtcatggaactatagcaagtgatctc 120

QY 122 ggtgagctgctgtggaacgaattccctccctagagtgccaatctttccattcaac 181

DB 121 ggaagactgctgtggaacgaattccctccctccctcggtgcgtgccaatctttccattcaac 180

6

QY 182 acctcagtcgtgtacaaaaagactctgtgtgtagaatccaggttccacctttcaacatc 241

DB 181 acctcagtcgtgtacaaaaagactctgtgtgtagaatccaggttccacctttcaacatc 240

QY 242 gctaaagccaaagccacctgtgagtgtctgtctaggtccctacatccagcttgagttat 301

DB 241 gctaaagccaaagccacctgtgagtgtctgtctaggtccctacatccagcttgagttat 300

QY 302 gatacagtggtcatctacctaagaacatggtctccctccctcagctcagcttctcagtt 361

DB 301 gatacagtggtcatctacctaagaacatggtctccctccctcagctcagcttctcagtt 360

QY 362 ggtgtatcctactgtgaaagctcttcgaggagctgtgatatgatatcagacagtcacaa 421

DB 361 ggtgtatcctactgtgaaagctcttcgaggagctgtgatatgatatcagacagtcacaa 420

QY 422 gagaagaagaatgataaagctctccctgtgtggaagccatacatatgtctgcaagctccg 481

DB 421 gagaagaagaatgataaagctctccctgtgtggaagccatacatatgtctgcaagctccg 480

QY 482 aaagaagaatgctccaatggtccttgaccctctgacctgtgcttaccctacatattcttcat 541

DB 481 aaagaagaatgctccaatggtccttgaccctctgacctgtgcttaccctacatattcttcat 540

QY 542 gtggaacctgtgtaaaagacttgaaattcagagctcatctggaagcctactagatgtaga 601

DB 541 gtggaacctgtgtaaaagacttgaaattcagagctcatctggaagcctactagatgtaga 600

QY 602 gggagctctggccaaagaaagacacagaccttgacaaatattacactcttctgtcgtta 661

DB 601 gggagctctggccaaagaaagacacagaccttgacaaatattacactcttctgtcgtta 660

QY 662 ttgatgaaggaaagtgtgacactcagaacaaagacctcctgtagcagatagagat 721

DB 661 ttgatgaaggaaagtgtgacactcagaacaaagacctcctgtagcagatagagat 720

QY 722 gctgcatctgctcggtcggtggtcctaaatgtcacacagtcacatgttatgttaaacag 781

DB 721 gctgcatctgctcggtcggtggtcctaaatgtcacacagtcacatgttatgttaaacag 780

QY 782 ctgcgaagctcgtatgtgatgacaggaatcagctatgtgagatgtgataagggc 841

DB 781 ctgcgaagctcgtatgtgatgacaggaatcagctatgtgagatgtgataagggc 840

QY 842 acctccttgaaagcactcaatatctcctcgaagtgcaacatctctgtggaagaaact 901

DB 841 acctccttgaaagcactcaatatctcctcgaagtgcaacatctctgtgtgaaacact 900

QY 902 cgcgaagcgtctctggaatctgcgcacataacttctcctactgtcgtcgaacactgtgat 961

DB 901 cgcgaagcgtctctggaatctgcgcacataacttctcctactgtcgtcgaacactgtgat 960

QY 962 gaccttggaacgtttctactcgttttgcataatcttcccaacaatatgtatgcatgaga 1021

DB 961 gaccttggaacgtttctactcgttttgcataatcttcccaacaatatgtatgcatgaga 1020

QY 1022 gcttatgtcaaaagttagacaagctgtccagaggaaccccaactacgaatgaaaaataatgaa 1081

DB 1021 gcttatgtcaaaagttagacaagctgtccagaggaaccccaactacgaatgaaaaataatgaa 1080

QY 1082 gaagcgaagaactatgatatgatacttactgatactctgaaatgtgagtgtcaggtttgat 1141

DB 1081 gaagcgaagaactatgatatgatacttactgatactctgaaatgtgagtgtcaggtttgat 1140

QY 1142 gatgaacactctcctctcttatccaattcgcgtcagttgccaagaagatccctaaact 1201

DB 1141 gatgaacactctcctctcttatccaattcgcgtcagttgccaagaagatccctaaact 1200

QY 1202 ttgggtacattacatctgtcgtgaaagaggaactgtggaactatcctccctagctccgcgc 1261

DB 1201 ttgggtacattacatctgtcgtgaaagaggaactgtggaactatcctccctagctccgcgc 1260

QY 1262 ccgagtacagaagaattataaaagtcaatatgtgaaacatggtccctcagcggattgtgtag 1321



|||||  
Db 1261 ccgatgacgaagttaataaagtaatttgacaatgcccctcagcgattggaag 1320  
QY 1322 aagtacaaaaaagtcgatattatggtcacaacgataaacccttaagactgtgaagct 1381  
Db 1321 aagtaaaaaaaagtcgatattatggtcacaacgataaacccttaagactgtgaagct 1380  
QY 1382 attcagatataacgaagaacttggagaccttcttataatggggaagtggagacacgt 1441  
Db 1381 attcagatataacgaagaacttggagaccttcttataatggggaagtggagacacgt 1440  
QY 1442 ttgatatattaagaatacgaagacagacacataacaatcctacccctcaggaatcact 1501  
Db 1441 ctcatattattaagaatacgaagacagacacataacaatcctacccctcaggaatcacc 1500  
QY 1502 gatgtccgtcccttgttatcttaagaagattacccaaaaggtataaacatttgaagattt 1561  
Db 1501 gatgtccgtcccttgttatcttaagaagattacccaaaaggtataaacatttgaagattt 1560  
QY 1562 ccaattctgcaggaagaataatccaataataaattgacagtgactgtagaagaatggcca 1621  
Db 1561 ccaattctgcaggaagaataatccaataataaattgacagtgactgtagaagaatggcca 1620  
QY 1622 actaaatcagatccctcgttgcctgacccgctattactagtcttgcgttaaatatgagaga 1681  
Db 1621 actaaatcagatccctcgttgcctgacccgctattactagtcttgcgttaaatatgagaga 1680  
QY 1682 gatctagcttcaagactcaatcttgaccttccctatctgctacaagaagatctgtatagata 1741  
Db 1681 gatctagcttcaagactcaatcttgaccttccctatctgctacaagaagatctgtatagata 1740  
QY 1742 agaggaaacacgaataatgtcagaagaagaatgtcaatccctgttcttcgtattgtatga 1801  
Db 1741 agaggaaacacgaataatgtcagaagaagaatgtcaatccctgttcttcgtattgtatga 1800  
QY 1802 aacccaagctgtgtaacctcacagaataatacaacgcttcttccccaatccagcttggagt 1861  
Db 1801 aacccaagctgtgtaacctcacagaataatacaacgcttcttccccaatccagcttggagt 1860  
QY 1862 caactgtgagatcccaagattcccaagctcccaacataatgacagcatcaatggtctatgt 1921  
Db 1861 caactgtgagatcccaagattcccaagctcccaacataatgacagcatcaatggtctatgt 1920  
QY 1922 ttgtatagtttcagttgctgtagttgtttgcagtgtgacatggtgacatggtatcttaagc 1981  
Db 1921 ttgtatagtttcagttgctgtagttgtttgcagtgtgacatggtgacatggtatcttaagc 1980  
QY 1982 attggaacacagactgactccttctgtcttcttcttctctgatatataccttcaaacaaa 2041  
Db 1981 attggaacacagactgactccttctgtcttcttcttcttctgatatataccttcaaacaaa 2040  
QY 2042 atgtgtctataagaacacacaccaccctattcccatctcagagagaatactgtctcattgtcg 2101  
Db 2041 atgtgtctataagaacacacaccaccctattcccatctcagagagaatactgtctcattgtcg 2100  
QY 2102 atgaaaaaacacgctatgactctggaggtgcacaacacacagacttctggaacagagc 2161  
Db 2101 atgaaaaaacacgctatgactctggaggtgcacaacacacagacttctggaacagagc 2160  
QY 2162 atgacgccttactgaaggtttctatgtgtgacaagaacacatggtgatatattacgaagac 2221  
Db 2161 atgacgccttactgaaggtttctatgtgtgacaagaacacatggtgatatattacgaagac 2220  
QY 2222 agttatgaagatatcttcagatatcttgcgtgagtataaaaaacatgccaattgacaagaagc 2281  
Db 2221 agttatgaagatatcttcagatatcttgcgtgagtataaaaaacatgccaattgacaagaagc 2280  
QY 2282 ttctcccaagattcaagacacccctagacactagcaaaaagaatttaatgacccccca 2341  
Db 2281 ttctcccaagattcaagacacccctagacactagcaaaaagaatttaatgacccccca 2340  
QY 2342 gtcttgaacgcacatcaacgggaataactcgttactacttctcagtcagatacaagaagaa 2401  
|||||

Db 2299 gtcttgaacgcacatcaacgggaataactcgttactacttctcagtcagatacaagaagaa 2358  
QY 2402 attgacatgatatatccatatagtttgaaatgaagaagaatttgcatttatgat 2461  
Db 2359 attgacatgatatatccatatagtttgaaatgaagaagaatttgcatttatgat 2418  
QY 2462 gagatgaanaatcagaagcccccgagcttcaaaaagaacacagacatttctatgtc 2521  
Db 2419 gagatgaanaatcagaagcccccgagcttcaaaaagaacacagacatttctatgtc 2478  
QY 2522 gcaatgagaagccttggagattatggagatgagctccacacatgttctaagaagaag 2581  
Db 2479 gcaatgagaagccttggagattatggagatgagctccacacatgttctaagaagaag 2538  
QY 2582 gctcagaatgagcgttccctcagttcaagaagtttcttccaagaatttactgtaggc 2641  
Db 2539 gctcagaatgagcgttccctcagttcaagaagtttcttccaagaatttactgtaggc 2598  
QY 2642 tcccttactcagccctatataccgttggagaactaaatgaacatttggagctcctggggcca 2701  
Db 2599 tcccttactcagccctatataccgttggagaactaaatgaacatttggagctcctggggcca 2658  
QY 2702 tatataagacagaagttgagataataatcaatggttaacttccagaatcagccctcgt 2761  
Db 2659 tatataagacagaagttgagataataatcaatggttacttccagaatcagccctcgt 2718  
QY 2762 ccctattcccttactatcttaacttcttcttataagagaatacagaaggcagaagcaaa 2821  
Db 2719 ccctattcccttactatcttcttccctcaatatacatalgaggaagatcagaaggcagaagaa 2778  
QY 2822 cctagaanaaacttgtcagaagcctaaatgaacacaaacttacttcttgaagaatgcaaat 2881  
Db 2779 cctagaanaaacttgtcagaagcctaaatgaacacaaacttacttcttgaagaatgcaaat 2838  
QY 2882 cataatgacccacataaagaatgagattgacatgcaaaagcctggacttctctatgtat 2941  
Db 2839 cataatgacccacataaagaatgagattgacatgcaaaagcctggacttctctcagatgtc 2898  
QY 2942 gaccttgaaaagaatgtgcactagagcctgattggagcccttctgtctgcacacatac 3001  
Db 2899 gaccttgaaaagaatgtgcactagagcctgattggagcccttctgtctgcacacatac 2958  
QY 3002 acactgaacctgtcatgtagagacaagtacacagatcttctcgtttttcac 3061  
Db 2959 acactgaacctgtcatgtagagacaagtacacagatcttctcgtttttcac 3018  
QY 3062 atcttctgataagaccaaagctggtacttcaactgaanaataatggaagaacatgcaaggtc 3121  
Db 3019 atcttctgataagaccaaagctggtacttcaactgaanaataatggaagaacatgcaaggtc 3078  
QY 3122 ccccttgcataatccagattggaagttcccaactttaaagaagatatatgcttccatgtcaatc 3181  
Db 3079 ccccttgcataatccagattggaagttcccaactttaaagaagatatatgcttccatgtcaatc 3138  
QY 3182 aatggtctacataatgataacacactggtctgtagtaatggtcgaagatacaagaatgcga 3241  
Db 3139 aatggtctacataatgataacacactggtctgtagtaatggtcgaagatacaagaatgcga 3198  
QY 3242 tggatctgcctcagcatgggcagcaaatgaaacatccattatccattcagtggaacat 3301  
Db 3199 tggatctgcctcagcatgggcagcaaatgaaacatccattatccattcctccggaacat 3358  
QY 3302 gtgttcaactgtacgaaaaaagaaggtataaaaatggcacatgtaacaatcctatccaggt 3361  
Db 3259 gtgttcaactgtacgaaaaaagaaggtataaaaatggcacatgtaacaatcctatcccgga 3318  
QY 3362 gtcttgaagacagtggaaatgttataccatccaagctggaatcttggcgggtgtagatgcct 3421  
Db 3319 gtcttgaagacagtggaaatgttataccatccaagctggaatcttggcgggtgtagatgcct 3378  
QY 3422 attgacgagacatcatatgctgtgagatgagcaacttcttctgtgtatcagaacaaatgtgt 3481  
Db 3379 attgacgagacatcatatgctgtgagatgagcaacttcttctgtgtatcagaacaaatgtgt 3438  
|||||



QY	3482	cagactcccttgggaatgctctcttgacacataagagatttccagattacagcttcaagga	3541
Db	3439	cgactccctctgggaatgctctcttgacacataatagaatttccagattacagcttcaagga	3498
QY	3542	caataggagcagtggggcccccaagcttgccagagattatatttccgattcaataatggcc	3601
Db	3499	caataggagcagtggggcccccaagcttgccagagattatatttccgattcaataatggcc	3558
QY	3602	tggagaccacaaggagcccttctcttgattcaaggtggatctgttgcgaccaatgatatt	3661
Db	3559	tggagaccacaaggagcccttctcttgattcaaggtggatctgttgcgaccaatgatatt	3618
QY	3662	caagcgatcaagaagcccaagggtgcgccgtcagaagtcttccagccctatactccagttt	3721
Db	3619	caagcgatcaagaagcccaagggtgcgccgtcagaagtcttccagccctatactccaat	3678
QY	3722	atcatcatatagctcttgaatgggaagaagtggcagacttactcgaggaaatctccatgga	3781
Db	3679	atcatcatatagctcttgaatgggaagaagtggcagacttactcgaggaaatctccatgga	3738
QY	3782	accttaatgctctctcttggcnaatgttgatctacatctggaataaaacaaatatttcaac	3841
Db	3739	acctcatgctctcttctttggcaatgttgatctacatctggaataaaacaaatatttcaac	3798
QY	3842	ccttccaattatgtctctgatatacatctcgttttgacacccaactaatatagaatctgcagcaat	3901
Db	3799	ccttccaattatgtctctgatatacatctcgttttgacacccaactaatatagaatctgcagcaat	3858
QY	3902	cttcgcgaatggaagtcttgatggccttgatttaaatggttgccagatctgagcatgtggaaatggag	3961
Db	3859	cttcgcgaatggaagtcttgatggccttgatttaaatggttgccagatctgagcatgtggaaatggag	3918
QY	3962	agtaaaagcaataatcagatgacacagatattacgcttcatcttcaacttacaataatatttggcc	4021
Db	3919	agtaaaagcaataatcagatgacacagatattacgcttcatcttcaacttacaataatatttggcc	3978
QY	4022	acctgtgtctcttcaaaaagctcgtacattcaactccccaaggaggagtaatgctctgagacct	4081
Db	3979	acctgtgtctcttcaaaaagctcgtacattcaactccccaaggaggagtaatgctctgagacct	4038
QY	4082	cagggtgaataatccaaaagaagtgcctgcagaatggacttccagaagaacatgaaagtcaaca	4141
Db	4039	cagggtgaataatccaaaagaagtgcctgcagaatggacttccagaagaacatgaaagtcaaca	4098
QY	4142	ggagtaactactcaaggagtaataaactctctgcttaccagacatgtaatgtgaaggaattccctc	4201
Db	4099	ggagtaactactcaaggagtaataaactctctgcttaccctctatgtaacgtgaaggaattccctc	4158
QY	4202	atctccagcagtgtaagaatgagtcgaatctgaatctcttttttccaagaatgcaagaatgaag	4261
Db	4159	atctccagcagtgtaagaatgagtcgaatctgaatctcttttttccaagaatgcaagaatgaag	4218
QY	4262	gtttttcaggagaaatcaagaactccttcaacacctgtgttgaaactctcagaccccgcgtta	4321
Db	4219	gtttttcaggagaaatcaagaactccttcaacacctgtgttgaaactctcagaccccgcgtta	4278
QY	4322	ctgactcgtactaccttcgnaattcacccccagagatttgggtgcacacagattgcccttgaagatg	4381
Db	4279	ctcactcgtactaccttcgnaattcacccccagagatttgggtgcacacagattgcccttgaagatg	4338
QY	4382	gaggtctctggctgcgagagcacagagactctaactg	4416
Db	4339	gaggtctctggctgcgagagcacagagactctaactg	4373

XX	Human Factor VIII-C analog having exon 14 deleted.
DE	
XX	Human Factor VIII-C analog; exon deletion; coagulation disorders;
XX	
KW	haemophilia; ss.
RW	
XX	
PN	EP265778-A.
XX	
PD	04-MAY-1988.
XX	
PF	14-OCT-1987; 87EP-0115043.
XX	
PR	15-OCT-1986; 86US-0919153.
XX	
PA	(RORE ) RORER INT OVERSEAS.
XX	
PL	Salver N, Drohan W;
XX	
DR	WPI; 1988-120930/18.
XX	
PT	Human Factor VIII-C analogue free of other proteins - produced from
PT	recombinant DNA and used in the treatment of coagulation disorders.
XX	
PS	Claim 3; Page 13; 42pp; English.
XX	
CC	The genetically engineered analogue can provide a dependable and
CC	readily available therapeutic agent to be used in the treatment of
CC	haemophilia and coagulation disorders in humans or animals.
CC	See also AA061343-45.
XX	
SO	Sequence 4830 BP; 1392 A; 1091 C; 1058 G; 1289 T; 0 other;

Query Match	94.0%;	Score 4152;	DB 9;	Length 4830;
Best Local Similarity	95.0%;	Pred. No. 0;		
Matches 4881; Conservative	0;	Mismatches 35;	Indels 198;	Gaps 1

[illegible]



OY	542	gtgactcgtgtaaagagcttggaattcaggtccctattggagccctactgatagtatgtagaa	601
Db	541	gtgactcgtgtaaagagcttggaattcaggtccctattggagccctactgatagtatgtagaa	600
OY	602	ggagctctgcgaagggaagagacacagacctctgcacaaattatatactatttgcgtta	661
Db	601	ggagctctgcgaagggaagagacacagacctctgcacaaattatatactatttgcgtta	660
OY	662	tttgatgaagggaanaagtttggcactcagaaacaanaagaaacctcttgatgcagatlaggat	721
Db	661	tttgatgaagggaanaagtttggcactcagaaacaanaagaaacctcttgatgcagatlaggat	720
OY	722	gctgcatactgctcgggacctggtccctaataatgcacacagctcaatggtttatgtaaacagctc	781
Db	721	gctgcatactgctcgggacctggtccctaataatgcacacagctcaatggtttatgtaaacagctc	780
OY	782	ctgcgaagctctgattggaatgcacagaagaatcaagcttatcttgacatgatttggaaatgggc	841
Db	781	ctgcgaagctctgattggaatgcacagaagaatcaagcttatcttgacatgatttggaaatgggc	840
OY	842	aacacctctgaagtgcactcaataatcctcccgaaagtcacacacttctctgttagagacat	901
Db	841	aacacctctgaagtgcactcaataatcctcccgaaagtcacacacttctctgttagagacat	900
OY	902	cgccagagcgctcctcttggaatactcgcacaataactttccttactcgtcacaacctcttgatg	961
Db	901	cgccagagcgctcctcttggaatactcgcacaataactttccttactcgtcacaacctcttgatg	960
OY	962	gaccttggaaagtcttctactggtttgcaatctctctccacaacaatcgatgagcatggaa	1021
Db	961	gaccttggaaagtcttctactggtttgcaatctctctccacaacaatcgatgagcatggaa	1020
OY	1022	gcttatctcaaaatgagcaactctgccaagaggaaacccaactcgcgaatgaaataatgaa	1081
Db	1021	gcttatctcaaaatgagcaactctgccaagaggaaacccaactcgcgaatgaaataatgaa	1080
OY	1082	gaagcggagaactatgatactgatacttactgatactcgaatggaatgagtgtgcaggttgat	1141
Db	1081	gaagcggagaactatgatactgatacttactgatactcgaatggaatgagtgtgcaggttgat	1140
OY	1142	gatgacaacctctctcttactatccaaatttgctcaagttggccaagaagcatcctaaact	1201
Db	1141	gatgacaacctctctcttactatccaaatttgctcaagttggccaagaagcatcctaaact	1200
OY	1202	ttggtaacttacaatctgctgtctgaaggaggagctgagcaatgagcccttgcctgc	1261
Db	1201	ttggtaacttacaatctgctgtctgaaggaggagctgagcaatgagcccttgcctgc	1260
OY	1262	cccgatgcagaaagtataaaagtcaataatttgaacaatgagccctcagcggatttgtag	1321
Db	1261	cccgatgcagaaagtataaaagtcaataatttgaacaatgagccctcagcggatttgtag	1320
OY	1322	aagtaacaaaaaagtcogatttatggcatacacagatgaaaccttaagaacctcgfgaagct	1381
Db	1321	aagtaacaaaaaagtcogatttatggcatacacagatgaaaccttaagaacctcgfgaagct	1380
OY	1382	attcagaataatcagaagaactctgggaaccttactttagtggggaagtttggagacaacctg	1441
Db	1381	attcagaataatcagaagaactctgggaaccttactttagtggggaagtttggagacaacctg	1440
OY	1442	ttgatataatttaagaatcagaacagaacagacataataacatctacccctcagatcact	1501
Db	1441	ttgatataatttaagaatcagaacagaacagacataataacatctacccctcagatcact	1500
OY	1502	gagtgtccgtctcttgcattccaagggaattaccacaaaagtgtaaaacatttgaagatttt	1561
Db	1501	gagtgtccgtctcttgcattccaagggaattaccacaaaagtgtaaaacatttgaagatttt	1560
OY	1562	ccaattctgcagaggaataatccaattcaataatgaagcagctgactgttagaagatggggca	1621
Db	1561	ccaattctgcagaggaataatccaattcaataatgaagcagctgactgttagaagatggggca	1620
OY	1622	actaaatcagatcctcgtgtgcctgacccgctcttaactcagtttcgtttaaatatggagaa	1681

Db	1621	actaaataagatcctcctggtgcctcctgagccgctatacttacttacttgcgttatcttgaataatgtagaga	1680
Oy	1682	gatacagcttaaggacacatctgagccctccctcatctgtctacaagaatctgttagataca	1741
Db	1681	gatacagcttaaggacacatcttggccctcctccatctgtctacaagaatctgttagataca	1740
Oy	1742	agaggaacaacagataatgtcagacaagaagaaatgtcatctcgtttctctgtatattgtatg	1801
Db	1741	agaggaacaacagataatgtcagacaagaagaaatgtcatctcgtttctctgtatattgtatg	1800
Oy	1802	aaccgaagctgtgtacccccaagagaaatatacaacgctttctcccaatccagctgtagtg	1861
Db	1801	aaccgaagctgtgtacccccaagagaaatatacaacgctttctcccaatccagctgtagtg	1860
Oy	1862	cagctcttgagatccaagaggtcccaagcctcccaaaatcaatgcagaatcaatgtgcattgt	1921
Db	1861	cagctcttgagatccaagaggtcccaagcctcccaaaatcaatgcagaatcaatgtgcattgt	1920
Oy	1922	tttgatagtttgcagctgtgcagttgtgtttgtcagtaggtgtgcatactgtgtacaatctaagc	1981
Db	1921	tttgatattttgcagctgtgtcagttgtgtttgttcattaggtgtgcatactgtgtacaatctaagc	1980
Oy	1982	attggaagacagactgtgactctccttctcgtctctcctcgtgtatataccttcaacaacaa	2041
Db	1981	attggaagacagactgtgactctccttctcgtctctcctcgtgtatataccttcaacaacaa	2040
Oy	2042	atgtgtctatgtagaagacaacacacacccattcccaattctctgagagaacgtctctatgtcg	2101
Db	2041	atgtgtctatgtagaagacaacacacccattcccaattctctgagagaacgtctctatgtcg	2100
Oy	2102	atggaaaaacccaagatctcatgtagatctgtgggtgtgcacaacatccagacttctcgaaacagagc	2161
Db	2101	atggaaaaacccaagatctcatgtagatctgtgggtgtgcacaacatccagacttctcgaaacagagc	2160
Oy	2162	atgagcgccttaactgtgaaggtttctagtgtgtgcagaagacaactgtgtattatacgaagac	2221
Db	2161	atgagcgccttaactgtgaaggtttctagtgtgtgcagaagacaactgtgtattatacgaagac	2220
Oy	2222	agttatgtagaagatatttcaagataactgtgtgtgagaanaaataatgtccattgtgaaccaaagagc	2281
Db	2221	agttatgtagaagatatttcaagataactgtgtgtgagaanaaataatgtccattgtgaaccaaagagc	2280
Oy	2282	tctctccagaattca-----	2296
Db	2281	tctctccagaattcaactattgtgactccttgccttgccttggataaacactatgttactcagata	2340
Oy	2297	-----	2296
Db	2341	ccaaaagaagatgtgaaatcccaagaagatcaccagaaaaaacagcttttaagaanaag	2400
Oy	2297	-----	2296
Db	2401	gataccatttgtccctgtaacgcgtgtgtgaagaacatcatgtacatgcagcaataatgtag	2460
Oy	2297	-----	2296
Db	2461	ggacaaaataagcccggaataagaagatcaacctgtgggcacgcaagatcagaatcgtgaagaagctg	2520
Oy	2324	tttaatgtccaccccaacagctctgtgaacgcgcatacgaatcgaaggaataactcgtactactct	2383
Db	2521	tgctctcaaaaacccaacagctctgtgaacgcgcatacgaatcgaaggaataactcgtactactct	2580
Oy	2384	cagtcacagatcaagaaggaattgtactatgtatgtatcacatatcagttgtgaatgaagaagaa	2443
Db	2581	cagtcacagatcaagaaggaattgtactatgtatgtatcacatatcagttgtgaatgaagaagaa	2640
Oy	2444	gatttgacatttatgtatgtgggagatgaanaatcgagaccccgagcttccaaagaanaaca	2503
Db	2641	gatttgacatttatgtatgtgggagatgaanaatcgagaccccgagcttccaaagaanaaca	2700
Oy	2504	cgacacatttattatgtctcagctgagagagcctctggaattatcggatgtagtagcttccca	2563







GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 18, 2001, 01:19:45 ; Search time 22045.3 Seconds  
(without alignments)  
3099.125 Million cell updates/sec

Title: US-09-689-430-1\_COPY\_419\_4835  
Perfect score: 4417  
Sequence: 1 catgaatagactctcca.....agcacagagactctactga 4417

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:\*  
1: gb\_ba1:\*  
2: gb\_ba2:\*  
3: gb\_ba3:\*  
4: gb\_in1:\*  
5: gb\_in2:\*  
6: gb\_in3:\*  
7: gb\_om:\*  
8: gb\_ov:\*  
9: gb\_pat1:\*  
10: gb\_pat2:\*  
11: gb\_ph:\*  
12: gb\_pil:\*  
13: gb\_p12:\*  
14: gb\_p13:\*  
15: gb\_p14:\*  
16: em\_ba1:\*  
17: em\_ba2:\*  
18: em\_fun:\*  
19: em\_htgo\_hum:\*  
20: em\_htgo\_inv:\*  
21: em\_htgo\_rod:\*  
22: em\_htg\_hum1:\*  
23: em\_htg\_hum2:\*  
24: em\_htg\_hum3:\*  
25: em\_htg\_hum4:\*  
26: em\_htg\_hum5:\*  
27: em\_htg\_hum6:\*  
28: em\_htg\_hum7:\*  
29: em\_htg\_hum8:\*  
30: em\_htg\_inv1:\*  
31: em\_htg\_inv2:\*  
32: em\_htg\_other:\*  
33: em\_htg\_rod:\*  
34: em\_hum1:\*  
35: em\_hum2:\*  
36: em\_hum3:\*  
37: em\_hum4:\*  
38: em\_hum5:\*  
39: em\_hum6:\*  
40: em\_hum7:\*  
41: em\_in:\*  
42: em\_om:\*  
43: em\_or:\*

44: em\_ov:\*  
45: em\_pat:\*  
46: em\_ph:\*  
47: em\_pl:\*  
48: em\_ro:\*  
49: em\_sts:\*  
50: em\_sy:\*  
51: em\_un:\*  
52: em\_v1:\*  
53: gb\_sts1:\*  
54: gb\_sts2:\*  
55: gb\_sts3:\*  
56: gb\_sy:\*  
57: gb\_un:\*  
58: gb\_v11:\*  
59: gb\_v12:\*  
60: gb\_htg1:\*  
61: gb\_htg2:\*  
62: gb\_htg3:\*  
63: gb\_htg4:\*  
64: gb\_htg5:\*  
65: gb\_htg6:\*  
66: gb\_htg7:\*  
67: gb\_htg8:\*  
68: gb\_htg9:\*  
69: gb\_htg10:\*  
70: gb\_htg11:\*  
71: gb\_htg12:\*  
72: gb\_htg13:\*  
73: gb\_htg14:\*  
74: gb\_htg15:\*  
75: gb\_htg16:\*  
76: gb\_htg17:\*  
77: gb\_htg18:\*  
78: gb\_htg19:\*  
79: gb\_htg20:\*  
80: gb\_htg21:\*  
81: gb\_htg22:\*  
82: gb\_htg23:\*  
83: gb\_htg24:\*  
84: gb\_htg25:\*  
85: gb\_pr1:\*  
86: gb\_pr2:\*  
87: gb\_pr3:\*  
88: gb\_pr4:\*  
89: gb\_pr5:\*  
90: gb\_pr6:\*  
91: gb\_pr7:\*  
92: gb\_pr8:\*  
93: gb\_pr9:\*  
94: gb\_ro1:\*  
95: gb\_ro2:\*  
96: gb\_in4:\*  
97: gb\_pr10:\*  
98: em\_ba3:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4407.4	99.8	9354	9 AR003585	AR003585 Sequence
2	4357.4	98.7	4670	9 AR110040	AR110040 Sequence
3	4148.8	93.9	4278	10 I08644	I08644 Sequence 4
4	4147.2	93.9	4281	10 I08643	I08643 Sequence 3
5	4085	92.5	4348	10 I08642	I08642 Sequence 2
6	4085	92.5	4551	10 I08641	I08641 Sequence 1
7	3743.8	84.8	5035	9 AR034084	AR034084 Sequence
8	3743.8	84.8	5035	9 AR071306	AR071306 Sequence



9	3189..8	72..2	4334	9	AR029098	Sequence
10	2603..8	58..9	4451	9	AR110041	Sequence
11	2336..4	52..9	9009	9	AR003710	Sequence
12	2336..4	52..9	9009	9	AR029065	Sequence
13	2336..4	52..9	9009	10	I31901	Sequence
14	2336..4	52..9	9009	10	I63424	Sequence
15	2336..4	52..9	9009	97	HUMFVIT1	Sequence
16	2336..4	52..9	7056	10	I27063	Sequence
17	2334..8	52..9	8241	9	A05328	Sequence
18	2334..8	52..9	8241	9	A07042	Sequence
19	2334..8	52..9	8831	10	E00527	Sequence
20	2334..8	52..9	9029	97	HUMFVIT1	Sequence
21	2333..8	52..8	7272	10	I05404	Sequence
22	2333..2	52..8	8967	9	AX052730	Sequence
23	2333..2	52..8	8967	10	I71409	Sequence
24	2333..2	52..8	8967	93	HSFVIT1R	Sequence
25	2328..4	52..7	7440	10	I08345	Sequence
26	2314	52..4	7440	10	I08457	Sequence
27	2275..2	51..5	6999	10	I77105	Sequence
28	1848	41..8	7032	7	AF016234	Sequence
29	1840	41..7	7145	7	AF049489	Sequence
30	1801..8	40..8	7493	9	AR003712	Sequence
31	1801..8	40..8	7493	9	AR029067	Sequence
32	1801..8	40..8	7493	10	I63427	Sequence
33	1801..8	40..8	7493	94	MUSCCVIT1	Sequence
34	1752..2	39..7	6539	7	SSU49517	Sequence
35	1751..2	39..6	6402	9	AR029097	Sequence
36	1661	37..6	1993	10	I02047	Sequence
37	1603..2	36..3	1728	10	I02054	Sequence
38	1240..4	28..1	1623	10	I27064	Sequence
39	991..4	22..4	3852	10	I04400	Sequence
40	990..8	22..4	3852	10	E00422	Sequence
41	893..4	20..2	1130	9	AR003711	Sequence
42	893..4	20..2	1130	9	AR029066	Sequence
43	893..4	20..2	1130	10	I31900	Sequence
44	893..4	20..2	1130	10	I63426	Sequence
45	675..4	15..3	935	10	E00526	Part of PES

## ALIGNMENTS

RESULT 1  
LOCUS AR003585 9354 bp DNA  
DEFINITION Sequence from patent US 5744326.  
ACCESSION AR003585  
VERSION AR003585.1 GI:3964844  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 9354)  
AUTHORS Ill.C.R. and Bidlingmaier.S.  
TITLE Use of viral C15-acting post-transcriptional regulatory sequences  
to increase expression of intronless genes containing  
near-consensus splice sites  
JOURNAL Patent: US 5744326-A 2-28-Apr-1998;  
FEATURES  
source 1. 9354  
location/Qualifiers  
BASE COUNT 2506 a 2239 c 2161 g 2448 t  
ORIGIN

Query Match 99.8%; Score 4407.4; DB 9; Length 9354;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 4411; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 catgcaaatagagctccaccctgcttctctgctgcttggcattgccttaagtc 60  
DB 2964 CATGCTTATGAGCTCTCCACCTGCTCTTCTGTGCTTGGCATTCGCTTATGTC 3023

QY 61 caccagaataactacccctgggtgagtggaactgtcatggaactatatacgaaagtact 120  
DB 3024 CACCAGAGATACCTACCTGGGTCAGTGAACCTGCACTGATATGCAAAAGTATCT 3083  
QY 121 cggtagagctgctgtggagcgaagaattccctccctagagtgccaatcttccatcaa 180  
DB 3084 CGGTAGGTGCTGGAGCGCAAGATTTCTCTAGATGCGCAAAATCTTTCCATTCAA 3143  
QY 181 caccacagctgtgatacaaaaagactctgttgtagaatccaggttcaaccttcaacat 240  
DB 3144 CACCACATCGGTATCAAAAAGACTCTGTTGTAATATCAAGGTTTACCTTTCAACAT 3203  
QY 241 cgttaagcgaagccacccttgatggtgtctgtagtctcaatccatcaagctgagttta 300  
DB 3204 CGCTAAGCAAGGACACCTGATGGTCTGATGATCTTCAACAGGCTGAGCTTTTA 3263  
QY 301 tgatacaatggttcatataacttaagaataatggttccatcttctgtagtctttagtct 360  
DB 3284 TGATCAAGTGTGATTAACATTAAGAACATGGCTTCCATCTTCAGTTCATGCTGT 3323  
QY 361 tgggtatccctactggaagactctggaagagctggaatatagtatgaagacgaag 420  
DB 3324 TGCTATCTACTACTGGAAGACTTCTGAGAGGAGCTGAATATGATATGACACACTCAAG 3383  
QY 421 ggaagaagaatgataaagcttccctggtggaagcacaatataatgctggaagctct 480  
DB 3384 GGAGAAAGATGATTAAGTCTTCCCTGGTGAAGCATTATGATGCTGGCAGGCTCT 3443  
QY 481 gaaagaagaatgctcaatagctctgaccacactgtgcttccatccatatactttca 540  
DB 3444 GAAAGAGATGCTCAAAAGGCTCTGACCACTGCTTACCTTACCTATATCTTTCTCA 3503  
QY 541 tgtgacactgtgaaagacttgaattcagacctaatgtgaagccctactagatgtaaga 600  
DB 3504 TGTGACCTGTGAAAGACTTGAATTCAGGCTCATTTGAGACCTTACTAGATGTAAGA 3563  
QY 601 aggaagctgtgccaaggaagaacacagaactgtgcaaaattatactacttctgtct 660  
DB 3564 AGGAGTCTGGCCAAAGAAAGACACAGACCTTGCACAAATTTATCTATCTTTGCTGT 3623  
QY 661 attgataagaagaaagcttgcaactcagaagaacaaagaccccttgtagcagaatga 720  
DB 3624 ATTGATGAAGGAAAGATGCTGACCTGAGAACAAAGACTCTTGTGACGGATGAGGA 3683  
QY 721 tgcctacatgctcctgagccttgccctaaatgcaacagctcaatggtatgtaaacagtc 780  
DB 3684 TGCTCATCTGCTGGGCTGCTGCTTAAATGACACAGTCAATGTTATGTAACAGCTC 3743  
QY 781 tctgcaagctgcatgtagatgtagcacaaggaatcagctctatgtgcatgtagatggg 840  
DB 3744 TCTGCCAGGTCTGATTTGATGCCACAGGAATTCAGTCTATTTGGCATGTGGAATGGG 3803  
QY 841 caccacctcgaagtgaactcaatatctctgaaaggtacacattcttgtagaagaaca 900  
DB 3804 CACCACCTCTGAAGTGCCTCAATATTTCTGAAAGTACACATTTCTTGAAGAACCA 3863  
QY 901 tcgcaagcgtcctctggaatctgcgcaataacttcccttacttctgctcaaacacttgat 960  
DB 3864 TCGCCAGGCGTCTGGAATCTGCGCAATTAATCTTCTTCTGCTCAAAACACTTTGAT 3923  
QY 961 ggaacttggaagcttctactggttctgcatatctctccacaacaatgtagcatgga 1020  
DB 3924 GGACCTTGACAGTTTCTACTGTTTTCATATCTCTTCCACCAACATGATGACATGGA 3983  
QY 1021 agcttatgtcaaatgtagcaagctgtccagagaagaccccaactacgaatgaaataatga 1080  
DB 3984 AGCTTATGTCAAAAGTAGACACTGTCCAGAGAACCCCAATAGCAATGAATAATATGA 4043  
QY 1081 aagaacggaagactatgtagatctactacatgcatctgaaatgtagtgcaggttga 1140  
DB 4044 AGAAGCGAAGACTATGATGATGATCTTACTGATTTCAATATGATGTGTGACAGTTTGA 4103  
QY 1141 tgatgacaactctcctcttaltccaaatcgtcaggtgccaagaagatcctaaac 1200



Db 4104 TGAAGACAACTCTCTTCTTATCCAAATTCGCTCAGTTGCCAAGAAGCATCTTAAAC 4163  
Qy 1201 ttgggtataactatgtctgtcgtgaagagagactgggactatgctcccttagtccctgc 1260  
Db 4164 TTGGGTATATTACATTCGTCTGTAAGAGAGAGACTGGACATATCTCTCTAGTCCCTGC 4223  
Qy 1261 ccccgatbagaagaagttataaagtcfaatlttgacaatgagccctcagggatgtgtag 1320  
Db 4224 CCCCAGATGACGAAGGTTTATTAAGTCATTTTAAACATGCCCCCTCAGCGGATGGTAG 4263  
Qy 1321 gaagtacaaaaaagtcgcatltatgcatatacagatgaaaccttgaagctgtgaagc 1380  
Db 4384 GAAGACAAAAAAGTCGATTTATGGCATACAGATGAACCTTTAAGACTCGTGAAGC 4343  
Qy 1381 tattcagcatgaaatcagaatccttgggaaccttactcttaagggaagttggaacaact 1440  
Db 4344 TATTTCACATGAATCAGAAATCTTGGAGCTTTACTTTATGGGAGAGTTGAGACACT 4403  
Qy 1441 gtgattatatattaagaatcagaacagacacataacatcfaacctcaggaatcac 1500  
Db 4404 GTTATTTATTTTAAATCAGACAGACACATATACCTTACCTCAGGAAATCAC 4463  
Qy 1501 tgatgtccgtcccttgtatccaagagatliaccaaaagtgtaaaacatttgaagattt 1560  
Db 4464 TGATGTCCGTCTTGTATTCAGAGAGATTCACAAAAGGTGTAAACATTTGAAGGATTT 4523  
Qy 1561 tcgaattctgcagagagaataatacaataatgacagtgacgtgagaagatggcc 1620  
Db 4524 TCCATTTCTGCAGAGAAATATTCAAATATTAATGACATGACTGTAGAAATGGGCC 4583  
Qy 1621 aactaaatcagaatcctcgtgagccctgacccctactactcagttctgttaatataggag 1680  
Db 4584 AACTAATTCAGATCTCTGGTCCCTGACCCGCTATTACTCTAGTTTCTGTTAATGAGAG 4643  
Qy 1681 agatcagctcagagactatggccctctccatctgcatctgtacaaaagatctgtagtca 1740  
Db 4644 AGATCTGAGCTTCAGAGACTCATTTGGCCCTCTCCATCTGTACAAAAGATCTGTAGATCA 4703  
Qy 1741 aagggagaaacagtaatgtcagaagaaggaatgatcactcgttcttctgtatgtga 1800  
Db 4704 AAGGGAAACCCAGATATGTGCAGCAAGAGAAATGTATCTCTGTATTTGATGA 4763  
Qy 1801 gaaccgaagctgtgacctcacaagaatatacaacgcttctcccaatccagctgagtc 1860  
Db 4764 GAACCGAAGCTGTACTCTCAGAGATATATACAAGGCTTTCTCCCAATCCAGCTGGAGT 4823  
Qy 1861 gcagcttgagatccagaagttccaaagccctccaaatcatgacagacatcaatgtgctatgt 1920  
Db 4824 GCACCTTGAGGATCCAGAGTTCCAAAGCCTCCAAATCATGCACAGATCAATGCTATGT 4883  
Qy 1921 tttagatagtttagagttgtcagttgtttgcatgagtgagcatcactgtaactcctcaag 1980  
Db 4884 TTTTGTATGATTTGCAATTTGTGATTTGTTCATGAGGTGGCATATGATTTCTTAAG 4943  
Qy 1981 catggaagcagaagctgactccttctgtctctctcctcgtatataccttcaaacaa 2040  
Db 4944 CATTTGAGACACAGCTACTCTCTTCTGTCTTCTCTCTGTGATATACCTTCAAAACAA 5003  
Qy 2041 aatgtctatgaaagacacactacacctatcccatctcagggagaaactgltctcatgtc 2100  
Db 5004 AATGTCTATGAAGACACACTCACCTTATCCCATCTTCAGAGAAACTGTCATGATGC 5063  
Qy 2101 gatggaaaaacccaggtatgagttctgtgggtgccaacactcagacttctggaaacaag 2160  
Db 5064 GATGGAAAACCCAGGTTATGGATTTGGGGTGCCACAACCTTTCGGAACAAGG 5123  
Qy 2161 catgaccgcttactgaaggttctcagttgtgacaagaacaactggtattatcagagga 2220  
Db 5124 CATGACGCGCTTACTGAAGGTTTCTAGTTGTGCAAGAACAACCTGTGATATTTACGAGGA 5183  
Qy 2221 cagttatgaaagataatccagcatactgtctgagtaaaacaatgcatctgaaaccaag 2280

Db 5184 CAGTTATGAAAGATATTTTCAGCATCTTGCTGATTAATAAACAAATGCCATTGAACCAAGAG 5243  
Qy 2281 ctctcccaagatccaagacccttagactaggcaaaagaatttataatgcccacc 2340  
Db 5244 CTCTCCCAAGATTTCAACACACCCTAGCATGAGCAAAAAGCAATTTAATGTGCCACCCAC 5303  
Qy 2341 agcttgaacgcatacaagggaaataactcgttactcctctcagttcagatccaagaga 2400  
Db 5304 AGCTTGAAACGCCATCACGCGGAATATACCTGTACTCTTCTAGTCAGATTCAGAGGA 5363  
Qy 2401 aatgacbatgatacatacatalcagttgaaatgaaagaagaagatttgacattatga 2460  
Db 5364 AATGACATATGATGATACCATATCATGTTGAATGAAAGAGAAAGATTTTGACATTTATGA 5423  
Qy 2461 tgaagatgaaataatcagaagccccgcagcttccaagaagaacaagacatttattgtc 2520  
Db 5424 TGAGAGTGAATAATCAGAGCCCCGACGCTTTCAAAAGAAAACAGACATTTTATTTGTC 5483  
Qy 2521 tgcagtgagaaggtctcgggataatggaatgagtagtgcctcccatgttctaaagaacag 2580  
Db 5484 TGCAGTGGAGAGGCTCTGGGATTTATGGATGATGACTCCCCACATGTTCTTACGAAACAG 5543  
Qy 2581 ggtcagaagtgagcagtgctccctcagttcaagaagaagtgcttccaggaattactatgag 2640  
Db 5544 GGCTCAGAGTGGAGTGTCCCTCAGTTCAAGAAAGTTGTTTCCAGGAATTTACTGATGG 5603  
Qy 2641 ctctcttaactaagcccttatccgttggaagaactaaatgaaatltgggacctcgtgggccc 2700  
Db 5604 CTCTTTACTAGCCCTTATACCTGTGAGAACTTAATGAACATTTGGACATCTGGGGCC 5663  
Qy 2701 atataatgaagcagaagttgagaataatcatatgttaacttcaagaaatcagagctccg 2760  
Db 5664 ATATATTAAGACACAAGTGTGAAGTAAATATCATGTGTAATCTTCAGAAATCAGGCTCTCG 5723  
Qy 2761 tccctatcctctatctcagccttattctctatgagaagatcagaagcagaaga 2820  
Db 5724 TCCCTATTTCTTATCTTACTACCTTATTTATTTATGAGAAAGATCAGAGGCAAGGACGA 5783  
Qy 2821 accatgaaaaaacttgtcagccctaataagaaacaaactacttggaaagtgaca 2880  
Db 5784 ACCTAGAAAACCTTTGTCAAGCTTAATGAACCAAAACCTTCTTTGGAAAGTGCAACA 5843  
Qy 2881 tcaatgtgaccccaataaagaatgagttgagctgcaaaagccttgggcttattctctgagt 2940  
Db 5844 TCATATGGACACCCTAAAGATGAGTTGACTGCAAAAGCCTGGGCTTATTTCTCGATGAT 5903  
Qy 2941 tgacctggaaaaaagatgtgacccaagcctgataatggaccctctggtctgcacaactaa 3000  
Db 5904 TGACCTGGAAAAGATGTGCACCTAGGCTGATTTGGAACCCCTTGTGTGCTGCCACATAA 5963  
Qy 3001 cacaactgaacccctgctatcagggagaagaagtgacagtgacaagaatttctcgtgtttcac 3060  
Db 5964 CACACTGAACCTCTCTCATGGGAGACAAGTGACAGTACAGGAATTTCTCTGTTTTCAC 6023  
Qy 3061 catcttggatgagacaaaaactgttactcactcagaaateltgaaagaactgcagggc 3120  
Db 6024 CATTTTGTATGAGACCAAAACCTGTACTTCTACATGAAATATGGAAGAATCTCAGGCG 6083  
Qy 3121 tccctgcaatatccagatlggaagatcccaactttaaagagaatatcgtctcatalgcaat 3180  
Db 6084 TCCCTGCAATATCCAGATGGAAGATCCACCTTTTAAAGAGAAATATGTGCTCCATGCAAT 6143  
Qy 3181 caatggtctaaataatgagatcacactcgtgttagtaigtctagagatcaaaagattcg 3240  
Db 6144 CAATGGCTTACTAATATGATACACTGCTGCTAGTATGAGCTGAGATCAAAAGGATTCG 6203  
Qy 3241 atggtatctgtcacaagtgggcaagcaatgaaacatcatctctatctcagtgagca 3300  
Db 6204 ATGTATCTCTGCTGCAATGGGCAACCATGAAACATCATTTCTATTCATTTGAGTGA 6263  
Qy 3301 tgtgttcaactgtacgaaaaaagaagagatbaaaatggcactgtlacaaatctctacag 3360  
Db 6264 TGTGTCTACTGTACGAAAAAAGAGAGATATTAATATGGCACTGTACATCTCTATCCAGG 6323



3361 ttttttgaagacagtggaatgtlaccatccaaagctggaaatttggcgggtggaatgcct 3420  
|||||  
6324 TGTTTTGAAGACAGTGAATGTATACATCCAAAGCTGGAATTTGGCGGTGGAATGCCT 6383  
3421 taattggcgaacatccacatgctggagatgagcaacttcttcggtgtacagcaataatgt 3480  
|||||  
6384 TATTGGCAGCATCTACATGCTGGGATGAGCACACTTTTTCGTGTACAGCAATTAAGTG 6443  
3481 tcagactccctgggaatgagcttcgagacacatagagaatttcagattacagcttcagg 3540  
|||||  
6444 TCAGACTCCCTGGGAATGCTTCTTGACACATTTAGAGATTTTTCAGATTACAGCTTCAGG 6503  
3541 acaaatgagacagtgaggcccaaaagctggccagacttcatttcgcgaatcaatcaatgc 3600  
|||||  
6504 ACAATATGACAGTGGGCCCAAGCTGGCCAGACTTCAATTATTCGGATCAATCAATGC 6563  
3601 ctgagagcaacgaagacccttctcttgatcaaggtlgaatcgtgtggaaccaatgatatt 3660  
|||||  
6564 CTGGAGGCACCAAGAGCCCTTTCTTGGATCAAGGTGATCTGTGGCACCAATGATTAAT 6623  
3661 tcagggacatcaagaccagggtgctgcagagaattctccagccttcacatctcagatt 3720  
|||||  
6624 TCAGGGCATCAAGACCCAGGGTGGCCGTCAGAAATTCTCCAGCCTCTACATCTCTCAGTT 6683  
3721 tatcatcatgataagttcttgatgggaagaagtgagacacttaccaggaatctccactgg 3780  
|||||  
6684 TATCATCATGATATAGTCTTTGATGGAGAGAGTGGCAGACTTATTCAGGAATTTCCACTGG 6743  
3781 aaccttaatggtctctcttggaatgttgatcattcgtggaataaaacaaatatttttaa 3840  
|||||  
6744 AACCTTAATGCTCTTGTGGCAATGTGGATTCATCTGGGAATAAACAAATATTTTAA 6803  
3841 cccccaattatgtctgatatacatccgtttgacccaacttcattagatctgcagac 3900  
|||||  
6804 CCTTCATTAATTTGCTCATACATCCGTTTGCACCCAACTCAATTATAGCATTCGCGAGCAC 6863  
3901 tcttcgcaatgagatgagtgagtgatgatttaaatagttgcaagatgccaatggaaatgga 3960  
|||||  
6864 TCTTCGCAATGAGTGTGATGGGCTGTGATTTAATAGTTTCAGCATGCCATGCCATGGGAATGGA 6923  
3961 gattgaagaacatatacagatgacagatactctcattccactcttaccatattgtttgc 4020  
|||||  
6924 GAGTAAGCAATATCAGATGCAAGATTAATCTTCAATCTTACCAATATGTTTTGC 6983  
4021 caactggtctcttcaaaagctcgactccaactccaaggaaggaatgagcttcgagacc 4080  
|||||  
6984 CACCTGGCTCTCTTAAAGCTCGACTTCACCTCCAAAGGAGAGATTAATGCTTGAGACC 7043  
4081 tcaagtgaaataatccaaagaagtgctgcaagtgagacttcagaagaacaaatgaagtcac 4140  
|||||  
7044 TCAGGTGAATATCCAAAGAGTGGCTCAAGTGACTTCAGAAAGACAAATGAAGTTCAC 7103  
4141 agggatgaactactcaaggaagtaaaatctctgtttaccagcatgatatgaaagagttcct 4200  
|||||  
7104 AGGAGTAATCTACTAGGAGTAATAATCTCTGTACACAGCATGTATGTAAGAGAGTTCT 7163  
4201 caatccagcagtcgaagatggccacatcagtgactccttcttccaaatggcaaaatgaa 4260  
|||||  
7164 CATCTCCAGAGTCAAGATGGCCATCAGTGGACTCTCTTTTTCAGAAATGGCAAACTAA 7223  
4261 ggttttccaaggaataatcaagactcctcaacactgltgltgaaactccttagaccacglt 4320  
|||||  
7224 GGTTTTTCAGGGAATCAAGACTCCTTCAACACTGTGGAATCTCTAGACCCACCGTT 7283  
4321 actgactgcactactcgaattcaacccccagagttggtgtgacacagatgctccctagaggt 4380  
|||||  
7284 ACTACTGCTAACCTTCAATTCACCCACAGTTGGGTGACACAGATTGCCCTGAGGAT 7343  
4381 ggaagttctggcgtgagagcaagagacttactga 4417  
|||||  
7344 GGAAGTTCTGGGCTGCGAGGCACAGGACCTTACTGA 7380

RESULT 2  
AR110040 4670 bp DNA PAT 14-FEB-2001  
LOCUS AR110040  
DEFINITION Sequence 41 from patent US 6114148.  
ACCESSION AR110040  
VERSION AR110040.1 GI:12826316  
KEYWORDS  
SOURCE  
ORGANISM Unknown.  
REFERENCE  
1 (bases 1 to 4670)  
AUTHORS Seed,B. and Haas,J.  
TITLE High level expression of proteins  
JOURNAL Patent: US 6114148-A 41 05-SEP-2000;  
FEATURES  
LOCATION/Qualifiers  
SOURCE 1..4670 /organism="unknown"  
BASE COUNT 1325 a 1072 c 1036 g 1237 t  
ORIGIN  
Query Match 98.7% Score 4357.4; DB 9; Length 4670;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 4358; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
59 gccaccagaagaatactaccctggtgagtgagacgtcatgtggaactatataagatgat 118  
|||||  
84 GCCACCAGAAATGATACCTGCTGCTGAGTGCAGTGCATGTCATGAGCTATATGCAAAAGTAT 143  
119 ctcggtgagctgctgtggaagcaagattccctcctagatgagcaaaatctttccattc 178  
|||||  
144 CTCGGGAGACCTCCCTGTGACGACCAAGATTCTCCTCAAGTGCMAAATCTTTTCCATTTC 203  
179 aacactcagtcgtgtacaaaagaactcgtttgagaatccaggtcacacttttaac 238  
|||||  
204 AACACCTCAGTGTGTACAAAAGACTGTTGTGAAATTCAGGATCACGATCTTTCAC 263  
239 atcgcataagcaagccaccctgagatgggtctgtaggtccctacatccagcagagtt 298  
|||||  
264 ATTCGCTAACCAAGGCGACCCCTGGATGGGTGCTGCTGATGCTACCATTCAGCGTAGGTT 323  
299 taatgataagtgatcattacacttaagaacatggtctccatccctgtcatgctcattgct 358  
|||||  
324 TATGATACAGTGTGATTAACATTAAGAAACATGCGTCCATCCGATGCTCATGCT 383  
359 gttgtgtatccctacttgaaagctctcgaaggagtgatataatgatacagacatcaa 418  
|||||  
384 GTTGTGTATCTTACTGGAAGCTTCTGAGGAGCTGAATATGATGATCAGACAGTCAA 443  
419 agggagaagaagatgataaagttccctggtgtagaagccatataatgcttgcaggtc 478  
|||||  
444 AGGGAGAAAGAGATGATTAAGTCTCCCTGGTGAACCATTAACATATGCTTGAGAGTTC 503  
479 ctgaaagaagaatggtccaaatggtcctcgaacactggtccttaccatcatcttctc 538  
|||||  
504 CTGAAGAGGAATGTCATGCTGATGCTGACCCACAGTGTCTCTTACTCATATATCTTCT 563  
539 catgtggacctgtgaaagacttgaaatcagagcctcatctggagccctactcagtagtga 598  
|||||  
564 CATGTGACCTGTGTAAGACTTGAATTCAGGCTCATTTGAGGCCCTACTACTGTATGTA 623  
599 gaaaggagctctggccaaggaagaacacagacacttgacaaattatatactacttttgtc 658  
|||||  
624 GAAAGAGTCTGGCCCAAGGAAGAACACAGACACTTGCAAAATTTATATCTACTTTTGGCT 683  
659 gtaattgataaggaagaaagttggcactcagaacaagaactccttgatgacaggaatag 718  
|||||  
684 GTATTGTGATGAAGGAAAAAGTTGGCAGTCAAGAAAGAAAGAACTCTTGATGACAGATAG 743  
719 gatgtcatatcgtcctgggacctgacctaaatgacacagatcaatgltatgataaagg 778  
|||||  
744 GATGCTGATCTGCTGGGCTGGCTTAAATATGACACAGTCAATGATTATGTAACAGG 803



OY	779	tctctccaaagtcgtatgtga tgcacagagaatacga tctatgtgc atgtatgtgaatg	838
Db	804	TCCTGGCAGGTCTGATTGGATGCCACAGGAATACGTATTATTTGGCATGTATTGGAAATG	863
OY	839	ggacacactctcctaagatgacataattccctcgaagatgcacaaattctctgtgaagAAC	898
Db	864	GGCACCACTCTGTAAATGGCACTAATAATTCTCTGAAAGTGCACACTTTCTTGTAAGAAC	923
OY	899	catgcacagcgctctctgtgaatctgcaccaataacttctctactgc tcaacaactcttg	958
Db	924	CATGCCACAGGGTCCTTGGAATCTGCCAATAACTTTCCTACTGCTCAAACTGTCG	983
OY	959	atggaactctggacaagttctctactcgttttgtcatactctctccaccaactatgtggcatg	1018
Db	984	ATGGACCTTGAGACGTTTCTACTGTTTTGCAATCTCTCCACCAACATGATGGCATG	1043
OY	1019	gaagctcatgtccaagatagaacaagt tccacagagaa ccccaactcgaatgaanaaat	1078
Db	1044	GAACCTTATGTCAAAGTAAAGACACGCTGCCAGAGAAACCCCAACTGCAATGAAAATAT	1103
OY	1079	gaagaaacggaagactatgtatgtatctctactatgtctgaatg atgtatgtgtaagttc	1138
Db	1104	GAAGAAACGGAAAGCATATGTATGTATCTTACTGATTTCGAATGTGATGTGTCAGGTTT	1163
OY	1139	gatgatgaacaactctctctctcttaccaatctgcgaatctgaagttccaaagaacatccctaa	1198
Db	1164	GATATATACAACTCTCTCTCTTATTCCAAATTCGCTGCAAGTTGCCAAGACATCTTAA	1223
OY	1199	acttggttacattacatctgcctcgtcgtcgaagagsgagctgggaactatgctcccttagtcctc	1258
Db	1224	ACTTGGGTACATTACATTTGGTCTGTGAAGAGAGAGACTGGGACTATAGTCTCTTGATCTTC	1283
OY	1259	gcccccgatacagaagttataaagaatcatcttgaacaatgagccctcagcggatgtgt	1318
Db	1284	GCCTCCCATGACAGAACTTTAAATCTCAATTTTGAACAAATGACCCTCACGGATTTGCT	1343
OY	1319	aggaagatcaaaaaaagtlccgattctcgtatctcgtacatacacagatgaaacctttaagaactcgtgaa	1378
Db	1344	AGGAAGTACAAAAAAGTCGATTTATGGCATACACAGATGAACCTTTAAAGACTCGTAA	1403
OY	1379	gctatccaagatgaatccaaggaactcttgggaacttacttaltbgggaaagtttggagacaa	1438
Db	1404	GCTTTTCAGATTAATTCAGGAATCTTTGGGACCTTTATATGGGGAAGTTGGAGACACA	1463
OY	1439	ctgtgtgcttatcttaagaacataagaacagacacataatacatctaccctccagaagatc	1498
Db	1464	CTGTGTATTAATTTAAGATCAAGCAACACCAATATACATCTACCTCTACAGGAATC	1523
OY	1499	actgaatgccgcctctgttattccaaggaattaccaaaaagtgtaaaacatttgaagat	1558
Db	1524	ACTGATGTCCGTCCTTTGTATTTCAAAGAAATTACCAAAAAGGTGTAAACATTTCAAGGAT	1583
OY	1559	tttccaactctgcacagagaataattccaataataaigtgacagatgactcgtatgaagatggg	1618
Db	1584	TTTTCAATTCGCCAGAGAAATATTCAAATTAATAATGACAGTACTGTGAAGAATGGG	1643
OY	1619	ccaactaatccagatcctcgttgctgcgcgaaccgctatactcagtttgcgttaatatbgag	1678
Db	1644	CCAACCTAAATACAGATCCTCGTGCCTGACCCGCTATTACTAGTTTTCGTAATATGAGG	1703
OY	1679	aggaatctagcttcagagactcatctggccctctccctcatctcgtctaaagaatctgtatgat	1738
Db	1704	AGGATCTTACCTTCAGACATCATTTGGCCCTTCCTCATCTATCTGCTCAAAAAGATCGTATGAT	1763
OY	1739	caaaaggaacaacagataatgtctacagaacaaggaatgtatctcgtttcttctgtatcttgat	1798
Db	1764	CAAAAGGAAGAAACCAAGTAATAGTCAAGCAAGAGGAATGTATCTGTTTTCGTATTTTGAT	1823
OY	1799	gagaaacgaagctgtgtactcctcacagaaatatacaacgcttctctcccaatccagcttga	1858
Db	1824	GAGAAACCGAAGCTGTGTACCTCAACAGGAATATATCAACGCTTTCTCTCCCAATCCAGCTGGA	1883

OY	1859	gtgcagcttgaagatccaaaggttccaaagctccaaatcaatgcaagcatcaatgycat	1918
Db	1884	gtgcagcttgaagatccaaaggttccaaagctccaaatcaatgcaagcatcaatgycat	1943
OY	1919	gttttcaatgttgaaggttgtaagttgtttgcataagagtgcaactggttaacttca	1978
Db	1944	gttttcaatgttgaaggttgtaagttgtttgcataagagtgcaactggttaacttca	2003
OY	1979	agcatlgaagacagacatgcttcccttcgcttccttcctgcgatalaactccaaac	2038
Db	2004	agcatlgaagacagacatgcttcccttcgcttccttcctgcgatalaactccaaac	2063
OY	2039	aaaatggctatgaagaacacatcccatlccactlctcagsgaaactgcttcacg	2098
Db	2064	aaaatggctatgaagaacacatcccatlccactlctcagsgaaactgcttcacg	2123
OY	2099	tcgattgaaaaacccaaggttcattgatttctggggtgcacaactcgaactcttgaaacga	2158
Db	2124	tcgattgaaaaacccaaggttcattgatttctggggtgcacaactcgaactcttgaaacga	2183
OY	2159	ggcatgacgcgccttaccgaaggttctcagtttgtaacaagaacacggtgatattacag	2218
Db	2184	ggcatgacgcgccttaccgaaggttctcagtttgtaacaagaacacggtgatattacag	2243
OY	2219	gaacgtatgaagatatlccaagaatctgctgagttaaaaaaatgccaatgtaaccaaga	2278
Db	2244	gaacgtatgaagatatlccaagaatctgctgagttaaaaaaatgccaatgtaaccaaga	2303
OY	2279	agcttctcccaagaattcaaaacacacttagacacaggaacaaacaaattaaatgcaaccca	2338
Db	2304	agcttctcccaagaattcaaaacacacttagacacaggaacaaacaaattaaatgcaaccca	2363
OY	2339	ccagctcttgaagaacccaatcagsgaaataaactcgtacacacttccaagccaataag	2398
Db	2364	ccagctcttgaagaacccaatcagsgaaataaactcgtacacacttccaagccaataag	2423
OY	2399	gaactlgaactatgatactaacacatcagttgaatagaagaagaatttgacattat	2458
Db	2424	gaactlgaactatgatactaacacatcagttgaatagaagaagaatttgacattat	2483
OY	2459	gataagatatgaataatcagagccccgcgagcttccaagaacaaacgcgacatatttat	2518
Db	2484	gataagatatgaataatcagagccccgcgagcttccaagaacaaacgcgacatatttat	2543
OY	2519	gctcagcttgagagagcctctggaattatgagatagtagcttcccaaatgcttgaagac	2578
Db	2544	gctcagcttgagagagcctctggaattatgagatagtagcttcccaaatgcttgaagac	2603
OY	2579	aggagctcaagatgycagtgctccctcagttccaagaagaattgtttlccagaattactgat	2638
Db	2604	aggagctcaagatgycagtgctccctcagttccaagaagaattgtttlccagaattactgat	2663
OY	2639	ggctcccttactcagcccttatccgltgagaaactaaatgaacatttgggaactcttggg	2668
Db	2664	ggctcccttactcagcccttatccgltgagaaactaaatgaacatttgggaactcttggg	2723
OY	2699	ccatataaagaagaagaattggaataatatcatgttaacttcaagaatacaggccct	2758
Db	2724	ccatataaagaagaagaattggaataatatcatgttaacttcaagaatacaggccct	2783
OY	2759	cgctccatttccctctcatlctcagccttatcttctatlgagaagaatcagagaaggaacga	2818
Db	2784	cgctccatttccctctctcatlctcagccttatcttctatlgagaagaatcagagaaggaacga	2843
OY	2819	gaaccttgaaaaaaacttgcacagcctaatgaaaccaaacttacttlttgaaagtga	2878
Db	2844	gaaccttgaaaaaaacttgcacagcctaatgaaaccaaacttacttlttgaaagtga	2903
OY	2879	catataatgcaacccaataaagaatgagtttgaatgtaagacccaagccttggttatcttcgat	2938
Db	2904	catataatgcaacccaataaagaatgagtttgaatgtaagacccaagccttggttatcttcgat	2963
OY	2939	gttaccctggaaaaaagatctgtaactcgaagccatgtaagaccccttcgcttcgcaact	2998



|||||  
Db 2964 GTTGACCTGGAAGATGTCTCAGCTCAGCCCTGATTTGGACCCCTTGTGCTGCCACACT 3023  
OY 2999 aacacgactgaacctgctcctcagtgagagacaagtgacagtaacaggaattgctgtcttc 3058  
Db 3024 AACCACTGAACCTGCTCATGAGGAGACAGTACAGTACAGGAATTTGCTGTGTTTTC 3083  
OY 3059 accactcttgaatgaaccaaagctgctactcactgaaatataatgaagaaactgcag 3118  
Db 3084 ACCATCTTTGATGAGACAAAGCTGTACTTCACTGAAATATGGAAGAAGAACTGCAGG 3143  
OY 3119 gctccctcgcaatatccagatggaagatccacatttaagagaattatcgcttcatagca 3178  
Db 3144 GCTCCGCAATATCCAAATGGAAGATCCACTTTTAAAGAAATATTCGCTTCATGCA 3203  
OY 3179 atcaatgcaataatgaatgaatacactactgcttagtaatgagtcagagatcaagaagat 3238  
Db 3204 ATCAATGGCTACATTAATGATACACTACTGCTTAGTAATGCTCAGAGATCAAGATT 3263  
OY 3239 cgaatgatactgctcagcatgagcgaatgaacaatcactatcattcactgagtgga 3298  
Db 3264 CGATGATATCTGCTCAGATGGGCAAGATGAATGAATCAATTCATTTCATTTCAGTGA 3323  
OY 3299 catgtctcactgtaagaaaaaagagagataaaatgcaactgtaaatctatcca 3358  
Db 3324 CATGTCTTCACTTACGAAAAAGAGAGATATAAATGGCAGCTGTACATCTATGCA 3383  
OY 3359 ggtgttttgaagacagatggaatgtttacacacaaagcttggaatttgcggtggaatgc 3418  
Db 3384 GGTGTTTTGAGACAGTGAATGTTACCATCCAAAGCTGGATTTGGGGTGAATGC 3443  
OY 3419 ctatctggcagacatactacatgcttggaatgaacacatttctggtgtaacagacataag 3478  
Db 3444 CTATTTGGCGAGCATATCATATGCTGGGATGAGACACTTTTTCGTTGACGCAATTAAG 3503  
OY 3479 tctcagaactccctcgggaatgctctcggacacatagagatttcagattacagcttca 3538  
Db 3504 TGTGAGACTCCCTGGGATGGCTTGTGACACATTAAGATTTTCAATTAACACTTCA 3563  
OY 3539 ggaacaatgagacagtgagcccaaaagcttgacagactcattatccggatcaatcaat 3598  
Db 3564 GGACAAATATGACAGTGGGCCCCAAAGCTGGCAGACTTCATTATCCGGATCAATCAAT 3623  
OY 3599 gacctgagaccacaaagaccccttctctggaatcaagtgtgactgctgagccaaatgact 3658  
Db 3624 GCCGAGAGACCAAGAGGCCCTTTCTTGATCAAGGTGATCTGTGGCAACCAATGAT 3683  
OY 3659 attcagggcatcaagaccagagtgcccgatcagaagttctccagcttcaactcctcaag 3718  
Db 3684 ATTTCACGGCATCAAGACCCAGGTCCTGCAAGATCTCCAGCCTCTACATCTCAG 3743  
OY 3719 tttaacatcatgatagtctctgataatggaagatgagcagactatccggaatccact 3778  
Db 3744 TTATATCATCATGTATAGTCTTGATGGGAAGAGTGGCAGACTTATCGAGGAATTCCT 3803  
OY 3779 ggaacacttaatgctctctcttgcaatggtgataatcaltcgggataaaacacaaatltt 3838  
Db 3804 GGAACCTTAATGTCTCTTGTGGCAATGTGATTCATCTGGATTAACACAAATTTT 3863  
OY 3839 aaacctccaattatgctcgatataccgcttgcaaccaactcatatagcattcgagc 3898  
Db 3864 AACCTTCATTAATGTGCTGATACATCGTTTGCAACCCATTAATAGATTCGAGC 3923  
OY 3899 accttgcagatgagtgatgagtgctgataatagtttcagacatccatttggaaatg 3958  
Db 3924 ACTTTTGCATGAGTGAATGGCTGTGATTTAAATAGTTACAGATTCCTATGGGAATG 3983  
OY 3959 gagaagtaaacataatagatgacagacattactgcttcatcctaactataatgctt 4018  
Db 3984 GAGAGTAAGCAATATATGAGATGACAGATTAATCTGCTTACTTACCAATATGTTT 4043  
OY 4019 gccacgtgtctccttcaaaagctcgacttcaactccaaggagagtaatgctcgag 4078  
|||||

Db 4044 GCCACCTGCTCTCTTCAAAAGCTCGACTTACCTCCAGGAGGAGTAATGCCGTGAGA 4103  
OY 4079 cctcagtgtaataatccaaaagatggtcgaagtgaacttccagaagaacatgaagtc 4138  
Db 4104 CCTCAGGTGATTAATCCAAAAGAGTGGCTGCAAGTGCAGTCTCCAGAAACATTAAGTTC 4163  
OY 4139 acaaggaataactactcaaggagtaaaatcctctgcttaccagacatgatatgaaagatc 4198  
Db 4164 ACAGAGTAACACTCAGGAGGATTAATCTGCTTACACACATGATGTGAAGAGATTTC 4223  
OY 4199 ctcatctcagcagatgcaatgagcatatgagatgagctcttlttgaagtgaagata 4258  
Db 4224 CTCATCTCAGAGTCAAGATGCGCATGAGACTCTCTTTTTCAGAAATGGCAAGTA 4283  
OY 4259 aagtttctcagggaaatcaaacactctcaccactggtgtgaaactctcagaccacg 4318  
Db 4284 AAGTTTTTCAGAGGAATCAACACACCTTCACTGAGTGGTGAACCTCTAGAACCCACG 4343  
OY 4319 ttactgactcgctacactcctgaattccaccagaagttggtgacagatgcccgaag 4378  
Db 4344 TTACTGACTCGCTACTTCAATTCACCCCGAGATTGGTGACACAGATTGCCCTGAGG 4403  
OY 4379 atggaagttcgtggtcgtcgagagcagaagacctctactga 4417  
Db 4404 ATGAGGTTCTGGGCTGCGAGCACAGGACCTCTACTGA 4442  
  
RESULT 3  
108644 LOCUS 108644 4278 bp PAT 02-DEC-1994  
DEFINITION Sequence 4 from Patent WO 8800831.  
ACCESSION 108644  
VERSION 108644.1 GI:588650  
KEYWORDS  
SOURCE  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 4278)  
AUTHORS Pasek,M.P.  
TITLE DNA SEQUENCES CODING FOR MODIFIED FACTOR VIII:C AND MODIFIED FACTOR VIII:C-LIKE POLYPEPTIDES AND PROCESSES FOR PRODUCING THESE  
POLYPEPTIDES IN HIGH YIELDS  
JOURNAL Patent: WO 8800831-A 4 11-FEB-1988;  
FEATURES  
source 1..4278  
location/Qualifiers  
BASE COUNT 1244 a 942 c 946 g 1146 t  
ORIGIN  
  
Query Match 93.9%; Score 4148.8; DB 10; Length 4278;  
Best Local Similarity 97.9%; Pred. No. 0;  
Matches 4270; Conservative 0; Mismatches 2; Indels 90; Gaps 3;  
  
OY 59 gccaccagaagatactactcctggtgacgtggaactgtcaatggaactatgcaaatgact 118  
Db 1 GCCACCAAGAAATACTACTGCTGGTGAAGTGAACGTGCATGAGGACTATATGCAAAAGTAT 60  
OY 119 ctggtgagcgtgcctggtgagcagaatlttctctctagagtgtccaaaatcttttcaatc 178  
Db 61 CTCGGTGAAGTCCCTGTGTGAGCAAGATTTCTCTGAGAGTCCAAATCTTTTCCATTTC 120  
OY 179 aacacctcagtcgtgtacaaaagactctgtttgaaatcagaagttcaccctttcaac 238  
Db 121 AACACCTAGTCTGTACAAAAGACTCTGTTTGAATAATTCACAGGATCACTTTTTCAC 180  
OY 239 atcgctaaagccaaagccaccccttgatgtgtctgtaagtgctcaccatccagcagtgagtt 298  
Db 181 ATGCGTAAGCCAAAGGCCACCTGATGGGCTGCTAGTCTCTACCAATCCAGCGCTGAGGTT 240  
OY 299 tatgatagatgtgtcattacaaactaagaacatggtctccactcctgtcagttcactgct 358  
Db 241 TATGATACAGTGTGATTAATTAAGAACATGGCTTCCATCCTGCTGACGTCTTCAATGCT 300  
|||||



QY 359 gttgtatctacttgaaagcttcgaggagctgaatatgatacagaccgcaa 418  
|||||  
Db 301 GTTGTGTATCTACTGGAAGCTTCTGAGGAGCTGAATATGATGATCGACAGCA 360  
QY 419 agggagaaagagaatgaataagcttcctcgtgtgaaagccatacatatgttcgagctc 478  
|||||  
Db 361 AGGAGAAAGAAAGATGAATGAATGCTTCCCTGTGTGGAAGCCATATATGTCTGCAAGTc 420  
QY 479 ctgaagaagaatggtccaatgctctgagccactgtgcttactactcaatctctc 538  
|||||  
Db 421 CTGAAGAGAAATGCTCAATGCGCTTGACCCACTGTGCTTACTACTCATATCTTCT 480  
QY 539 catctgagccgtgaaagacttgatctcagccctcatcttgagccctactagatgaga 598  
|||||  
Db 481 CATGTGACCTGCTTAAGAACTTAATTCAGGCTCATTTGAGGCCCTACTAGATGTAGA 540  
QY 599 gaaggagctctgccaagaaagacacagaccccttgacaaatctactactcttctgct 658  
|||||  
Db 541 GAAGGAGTCTGGCCAAAGCAACACACCTTGACAAATTTATACTACTTTTGTCT 600  
QY 659 gtacttgatgaaggaaagttgagctcagaaacaaagaactccttgatgagagatag 718  
|||||  
Db 601 GTATTGTATGAAGGAAAGTTGGCACTCAGAAACAAAGAACTCTTGATGACAGATAGG 660  
QY 719 gatgtcatctgtctgggctcgtgacctaaatgacacagtcacatggttatgtaacag 778  
|||||  
Db 661 GATCTCATCTGCTCTCGGCTGGCTTAATAATGCAACAGCAATGATGTATGTAACAGG 720  
QY 779 tctctg---cgaagctcgatgtgatacgaagaatcagtcatacttgcatgtgattgga 835  
|||||  
Db 721 TCTGTGTACCAAGTCTGATTTGATGTCACAGAAATCAGTCTTTGGCATGTGATTTGGA 780  
QY 836 atggagcaccactcctgaaagtgactcaatatcctcgaaggtcacacattctgtgag 895  
|||||  
Db 781 ATGGGCACTCTCTGGAAGTCACTCAATATTCCTCGAAGTCAACATTTCTGTAGG 840  
QY 896 aacatctgcaagggtcctcttggaatatctgcacaaataacttctactgtctcaaacctc 955  
|||||  
Db 841 AACCATGGCAGGCGTCTTGGAATCTCGCAATATCTTCTACTGCTCAAAACCTC 900  
QY 956 ttgatgaccttggaagcttctactgttttgcatactctctccacaaatcagtcagc 1015  
|||||  
Db 901 TTGATGACCTTGAGAGTTTCTACTGTGTTTGTATATCTCTTCCACCAAGATGATGGC 960  
QY 1016 atggaagctatgtcaaatgagacagctgtccagaaggaaacccaactagaatgaaaaat 1075  
|||||  
Db 961 ATGGAACTTATGTCAAAAGTAGACAGCTGTCCAGAGGAACCCCAATACGAATGAAAAAT 1020  
QY 1076 aatgaagaacggaagactatgatgatcttacttgatctgaaatgaaatgtgtcag 1135  
|||||  
Db 1021 AATGAAGAGCGGAAGACTATGATGATGATCTTACTATTTCTGAATGAGATGTGAGG 1080  
QY 1136 ttgatgatgaacactcctctccttataccaatctgcctcagttgccaagaagcactct 1195  
|||||  
Db 1081 TTTGATGATGACAACCTCTCTCTTATCCAAATTCGTCAGTTGCAAGAAACATCTCT 1140  
QY 1196 aaaacttgggtacattacattgtctgtgaagaggagacttggagactatgctccctagtc 1255  
|||||  
Db 1141 AAAAATTGGGTACTATTCATTTGCTGTGAAGAGAGACTGGGACTATGCTCTCTAGTC 1200  
QY 1256 ctgcgccccgatatcagaagtctataaagfcaaatcttgaaatgacccctcagcgagt 1315  
|||||  
Db 1201 CTCGCCCCGATGACAAAGTATTAAGTCAATATTTGAACAATGGCCCTCACGGAAT 1260  
QY 1316 gttgagaagtacaaaagctccgaatttatgtcatacacagatgaactttaaagcct 1375  
|||||  
Db 1261 GGTGGAAGTACAAAAGTCCGATTTATGSCATACACAGATGAACCTTTAAGACCTG 1320  
QY 1376 gaagctatcagcatgaatcagaactcttgggaccttactttaaagggaagttgagac 1435  
|||||  
Db 1321 GAACCTATTGAGCATGATGATCAGGAATCTTGAGACCTTACTTTATGGGAAGTTGGAGAC 1380

QY 1436 aactgtgatatattatctaaagaatccaagcaagcagaccataatacatctaacctcagga 1495  
|||||  
Db 1381 AACTGTTGATTTATTTTAAGAAATCAAGACAGACCATATATACATCTACCTCACGGA 1440  
QY 1496 atcacgtatgtccgtccctcttgatctcaaggagatctaacaaaggtgttaaacatltgaa 1555  
|||||  
Db 1441 ATCAGTGATGTCCGCTCTTTGTATTCAAGGAGATTACCAAAAGGTGTAAACATTTGAAG 1500  
QY 1556 gatttccaattctgcagagagaataatccaataataatgagcagtgtactgtagaat 1615  
|||||  
Db 1501 GATTTTCCAATTTGCCAGGAGAAATATTCAAATATTAATGGAAGTGTAGAGAT 1560  
QY 1616 gggccaaactaaatcagaatccctcgtgtgctgaaccgcatctacttaactgttcaataag 1675  
|||||  
Db 1561 GGGCAACTAATATAGATCTCTGCTGCTGACCCGCTATTACTTACTTGTGTTAATAG 1620  
QY 1676 gaagagatctagcttcaaggactatgtgacctctccatctctgtcaaaagaatctgta 1735  
|||||  
Db 1621 GAGAGATCTAGCTTCAGGACTATTGGCCCTCTCCATCTGCTCAAAAGATCTGTA 1680  
QY 1736 gatcaaaagagaaacacagataatgtcaagaagaagaatgtcatcctgttctgtattc 1795  
|||||  
Db 1681 GATCAAAAGAGAAACAGATATGTCAAGACAAAGGAATGCATCTGTTTCTGTATTT 1740  
QY 1796 gatgaagaccgaagctgtgacctcacaagagaatatacaacgcttctcccaatccagct 1855  
|||||  
Db 1741 GATGAGAACCGAAGCTGTACTCTACAGAGATATACAGCTTTCCTCCCAATCCAGCT 1800  
QY 1856 ggaagtcagcttgagataccaagattccaagctcccaacatcatgacaagcatcaatgagc 1915  
|||||  
Db 1801 GAGATGAGCTTGAGATTCAGATTCAGATTCCAAGCCCTCAACATCATGACAGATCAATGGC 1860  
QY 1916 taagtcttgaatagtttgcagttgtcagttgttgcataaggtgtgcatactgttactc 1975  
|||||  
Db 1861 TATGTTTTGATTTGTGCAATGTGTCAATGTTGTTCATGATGAGGCGATCTGTATATT 1920  
QY 1976 ctaagcatggaagacagactgactcctctctgtctctctctctctgatabaacttcaaa 2035  
|||||  
Db 1921 CTAAAGCATTTGGAGCACAGACTGACTCTCTTCTGTCTCTTCTCTGTGATATACCTTAAA 1980  
QY 2036 cacaaaaatgttctatagaagacacactacccatcttccactctcaagagaacatgtctc 2095  
|||||  
Db 1981 CACAAATGCTATGTAAGACACACTCACCTATTTCCATTTCTCAGAGAAATGCTCTTC 2040  
QY 2096 atgtcgaatggaataacccaagttctatgtatctctggtgtgccaacactcagacttccggaac 2155  
|||||  
Db 2041 ATGTGATGGAAGAAACCCAGGCTATGATTTCTGTGGGTGCCAACACTCGACTTTCGGAAC 2100  
QY 2156 aagaagcatgaccgcttactcgaaggttctcagttgtgacaagaacacactggtgatctaac 2215  
|||||  
Db 2101 AGAGGATGACCGCTTACTGTAAGGTTTCTAGTTGTGACAAAGAACACTGGTGATTTATAC 2160  
QY 2216 gaagacagtatagaatalcttcagcatctctgtctgtaataaanaatgtccatgtgaacca 2275  
|||||  
Db 2161 GAGGACAGTTATGAAGATATTTCAGCATCTGCTGATTAAGAAACATGCGCATGAAACA 2220  
QY 2276 agaaagcttctcccaagaatccaagaacccctagcaactagcgaagaagaatctaagccacc 2335  
|||||  
Db 2221 AGA----- 2223  
QY 2336 ccaccagctctgaaagccatccaacggaataactcgtactactcttcagtcagatcaa 2395  
|||||  
Db 2224 -----GAATATCTGTCTACTCTTTCAGTCAATCA 2256  
QY 2396 gaagaaatgactatgatacacaatcagttgaatcagttgaatgaagaagaatlttgacat 2455  
|||||  
Db 2257 GAGGAATTTGACTATGATGATACATATCAGTTGAATGAAGAAAGATTTTGAACATT 2316  
QY 2456 tatgatgagatgaataatcagagcccccagcttctcaaaaagaagaacagacactatctc 2515  
|||||  
Db 2317 TATGATGAGATGAATAATCAGAGCCCGCAGCTTTTCAAAAAGAAACACACACATATTTT 2376  
QY 2516 atgtcgtcagtgagagagccttggaatctatgagatgtgtagctcccaatglttctaaga 2575



|||||  
Db 2377 ATTCTGACAGTGGAGAGCCTCGGGATTATGGAGTAGAGCTCCCCACATCTTCTAATA 2436  
2576 aacggggtcagagtgagtgagtgctccagtgctcaagaagaagtgcttccaggaatttact 2635  
|||||  
Db 2437 AACGAGGCTCAGAGTGGCAGTGTCCCTCACTTCAAGAAAGTTGTTTCCAGAAATTTACT 2496  
2636 gatgctccttactcaagcccttatacgttggagaactaaatgaacatttggagactccty 2695  
2497 GATGGCTCCTTACTCAGCCCTTATACGTGAGAAATGAATGAACATTTGGAGCTCCG 2556  
2696 gggccatatataagagcagaagtttgaagataatcatgtaacttccagaataagcc 2755  
2557 GGGGCATATATAGAGCGAAGTTGAAATATATCATGTGAACCTTTCAAGAAATCAGGCC 2616  
2756 tctgtcccttactccttacttactcagcttacttcttataagagaatcatagaaggcaaga 2815  
2617 TCTGTCCCTTCTCTTCTTATTTCTAGCCTTATTTCTTATGAGGAAGATCAGAGCAAGCA 2676  
2816 gcgaacctagaaaaaacttctgcaagcctaaatgaaccaaaacttacttctgaaagtg 2875  
2677 GCAGACCTAGAAAAAATTTGTCAAGCCTAATGAACCAAAACCTTACTTTTGGAAAGTG 2736  
2876 caacataatgagcccaactaaagaatgagttgactgcaagcctgggcttacttctc 2935  
2737 CACATCATATAGGACACCCATGAAGATGAGTTGACTGCAAGGCCCTGGCTTATTTCTCT 2796  
2936 gatgtgacccggaagaaagatgtagcactcaagccgtatctgagcccttctgtctgcaac 2995  
2797 GATTTGACCTGGAAAAAGATGTGCACCTCAGGCCCTGATTTGGACCCCTTGTGCTGCCAC 2856  
2996 actaacaactgaacctgtcactgagagagaagtgacagtaacaggaatttgcctgtt 3055  
2857 ACTAACACACTGAAACCCCTGCTCATGTGAGACAGAGTACAGAGATTTCTCTGTCT 2916  
3056 t---tcaacatcttctgtagagaccaaaagctggtacttccactgaaataatgtaaaagaac 3112  
2917 TTCTCAGCATCTTTGATGAGAACCAAAAGCTGTACTTCACTGAAATAATGAAAGAAAC 2976  
3113 tgcagagctccctgcaatccagatggaatggaatccactttaaagaagaatcatcgctc 3172  
2977 TGCAGGGCTCCCTGCAATATTCAGATGGAAGATCCCACTTTTAAAGAGAAATATGCTTC 3036  
3173 catgcaatcaatggtctacataatgatacactggtctgtaatggtctcagaatcaa 3232  
3037 CATGCATCATATGCTTACATATGATGATACACTGCTGCTTGAATGGCTCAGAGATCAA 3096  
3233 aggatctgagtgatctgtcagcaatggcagcaatgaaataatccatcttcttcaattc 3292  
3097 AGGATTCGATGGTATCTGCTCAGCATGGGCAAGCAATGAAGAACATCATTTCTATTTCTC 3156  
3293 agtgagacatggtctcaactgtacgaaaaaagagagtaaaatgacactgtacacatcc 3352  
3157 AGTGACATGTGTCTACGTACGAAAAAAGAGAGATTAATAATGGCATGTACATCTCC 3216  
3353 taccaggtgttctttagagacagtgaaatgtaaccatccaaaagctggaatttgcgagtg 3412  
3217 TATCCAGAGTGTGTTTGAAGACAGTGAATGTATCCATCCAAAGCTGGAATTTGGGGGAGT 3276  
3413 gaatgcttattgagcagcatctacatgctgagatgagcaacacttcttctgtgtacagc 3472  
3277 GAATGCCCTTATTTGGCAGCATCTACATCTGGGATGAGCACACTTTTCTGTGTACAGC 3336  
3473 aataagtgtagactcccttggaatgagcttctgagacatataagatcttcaattaca 3532  
3337 AATAAGTGTGAGACTCCCTCGGAATGGCTTCTGACACATTAAGATTTTCAGATTACA 3396  
3533 gctcagagacaataatgagcagtgggcccaagctgcccagactcatatccgagta 3592  
3397 GCTTCAGAGCAATATGACATGTGGGCCCAAGAGCTGCCAGACTTCATTTATCCGAGATCA 3456  
3593 atcaatgcttgagacacaaagagcccttcttctgagcaaggtgagatcgtgtgacaca 3652  
|||||

Db 3457 ATCAATGCCGTGAGACCAAGAGGCCCTTTCTTGGATCAAGTGTGATCTGTGGCACCA 3516  
Qy 3653 atgatattcaagcatcaagaccaggttgcccgctcagaagttctcagccctacatc 3712  
|||||  
Db 3517 ANGATTATTCAGGATCAAGACCCAGGGTCCCGTCGAAGTTCTCCAGCTCTTACATC 3576  
3713 tctcagttatcatcatglatagttctgagtggaagatgagcaacttccagaagaat 3772  
3577 TCTCAGTTATTCATCATGTATGATCTTGATGGGAAGAGTGGCAGACTTATCGAGAAAT 3636  
Qy 3773 tccactggaaccttaatgcttcttcttggcaatgtgataatcatcttggtataaaacacaa 3832  
3637 TCCACTGGAACCTTATATGCTCTTCTTGGCAATGTGATTCATCTGGAATTAACACCAAT 3696  
Qy 3833 attttaaccctccaatattgtctcgatatacgcgttgcaccccaactcatatagcat 3892  
3697 ATTTTAAACCTCCCAATTTATGCTGATACATCCGTTTGGACCCCACTCATTAAGATT 3756  
Db 3757 CCGACACTCTTCGATGAGTGTGAGTGGCTGTGATTTAAATAGTTCCAGCATCCATTG 3816  
Qy 3893 cgaagcactcttcgacatgagttgagtggtctgataatgaatgctgcaatgcaatg 3952  
3757 CCGACACTCTTCGATGAGTGTGAGTGGCTGTGATTTAAATAGTTCCAGCATCCATTG 3816  
Qy 3953 ggaatgagagtaagaacataatcagatgacagatctcttacttacttactcaat 4012  
3817 GGAATGAGAGTAAGCAATATCATATGCACAGATTACTGCTTATCTTACTTATCCAAAT 3876  
Qy 4013 atgttgcacactgtctcccttcaaaagctgacttcaaccctcaagagagagatgac 4072  
3877 ANGTTGGCACCTGCTGCTCTTCAAAAGCTGACTCACTCCCAAGGAGAGTAATGCC 3936  
Qy 4073 tggagacctcaagtgataatccaaagaatggtctgcaagtgacttccagaagaatg 4132  
3937 TGGAGACCTCAGGTGAATAATCCAAAGAGTGGCTGCAAGTGGACTTCCAGAAACCAATG 3996  
Qy 4133 aaagtcaacagatgtaactcactcagagagtaaaactctgcttactcaagatgtagaag 4192  
3997 AAAGTCACAGGAGTACTCTCAGGAGTAATAATCTGCTTACACAGATGTATGTAAAG 4056  
Qy 4193 gagttccctacatccagcagctcaagatgagcatcagtgagactccttcttccagaatg 4252  
4057 GAGTTCCTCATCTCCAGAGTCAAGATGCCATCGATGAGGACTCTTTTTCAGAAATGCC 4116  
Qy 4253 aaagtaaggttcttcaagggaaatcaagaactccttcaacactgtgtgtaactctagac 4312  
4117 AAAGTAAAGGTTTTTCAAGGAAATCAAGACCTTTCACACTGTGGTGAACCTCTAGAC 4176  
Db 4117 AAAGTAAAGGTTTTTCAAGGAAATCAAGACCTTTCACACTGTGGTGAACCTCTAGAC 4176  
Qy 4313 ccaacgttactgactcgtctactccttcaatcccccagaagttggtgacacagatggc 4372  
4177 CCACGTTACTGACTCGCTACTCTTGAATTCACCCCAAGATGTGGTGCACCAATTTGCC 4236  
Db 4373 ctgaggaatgaggttctggtctgagagcagagacactctac 4414  
Db 4237 CTGAGATGGAAGTCTGTGGCTGGAGGACACAGACCTCTAC 4278

RESULT 4  
LOCUS 108643 4281 bp  
DEFINITION Sequence 3 from Patent WO 8800831.  
ACCESSION 108643  
VERSION 108643.1 GI:588649  
KEYWORDS  
SOURCE  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 4281)  
AUTHORS Pasek,M.P.  
TITLE DNA SEQUENCES CODING FOR MODIFIED FACTOR VIII:C AND MODIFIED FACTOR VIII:C-LIKE POLYPEPTIDES AND PROCESSES FOR PRODUCING THESE POLYPEPTIDES IN HIGH YIELDS  
JOURNAL Patent: WO 8800831-A 3 11-FEB-1988;  
FEATURES Location/Qualifiers  
Source 1..4281



BASE COUNT 1245 a 943 c 946 g 1147 t  
ORIGIN

Query Match 93.9% Score 4147.2; DB 10; Length 4281;  
Best Local Similarity 97.9% Pred. No. 0;  
Matches 4269; Conservative 0; Mismatches 3; Indels 90; Gaps 3;

QY 59 gccacccagaataatactactctggtggtcaggtggaactgtcattggaagctatataatgcaagatg 118  
|||||  
DB 4 GCCACCGAAGAACTACTACCTGGGTGCACTGAGTCACTGTCATGGAGCTATATCCAAAGAT 63  
QY 119 ctggttagctgctctggaagcaagatttcctcttagaagtgcacaaatctttccattc 178  
|||||  
DB 64 CTCGGTAGCGCTCTGTGAGACGAAGATTCTCTCTAGAGTCCAAATCTTTCCATTTC 123  
QY 179 aacacctcagtcgtgtacaaagaactcgtttgttagaatcacggttcacctttcaac 238  
|||||  
DB 124 AACACTCAGTCGTGTACAAAGAAGCTCTGTTGTAGATTCACGATCACCTTTTCAAC 183  
QY 239 atgcctaaagcccaagccacctgtgagtctgtctaggtcttaccatcccaagcctgaagtt 298  
|||||  
DB 184 ATGCGTAAAGCCAAAGCCACCTGGATGGTCTGCTAGGCTCTTACCATCCAGGCTGAGGTT 243  
QY 299 tatatacagtggtcatctataccttaagaacatggtcttccactcctgtcaggtctcatgct 358  
|||||  
DB 244 TATGATACAGTGTCATTTACCTTAAGACATGAGCTTCCCTCTGTGAGTCTTTCACTGCT 303  
QY 359 gtgggtatctctactgtgaaagctctctgaaggagctgaataatgaatgcagaccagtcaa 418  
|||||  
DB 304 GTTGGTGTATCTACTGGAAGGCTTCTGAGGAGCTGAATGTATGATGACAGACCAAGTCAA 363  
QY 419 agggagaagaagatgtataagttctccctggtgtgaagccaataatgtctgtcaggtc 478  
|||||  
DB 364 AGGAGAAGAAAGATGATTAAGTCTCCCTGTGGAAGCCATACATATGCTGTGACAGTCC 423  
QY 479 ctgaaagaagaatgtgtcccaatggtcctctgacccactgtgcttactctactatctctct 538  
|||||  
DB 424 CTGAAAGAGAAATGTCCTCAATGGCTCTGACCCACTGTGCTTACTCTCATATCTTTCT 483  
QY 539 catgtgacctgtgtataaagaactgaatlcagggcctcatctgaagccctactagtaatgta 598  
|||||  
DB 484 CATGTGAGACCTGTTAAAGACTTAATTCAGGCTCATTTGGAGCCCTACTAGTATGTAGA 543  
QY 599 gaaggaggtctgtgccaagaaagaacacagacctgtgcaaaatttatactacttttgc 658  
|||||  
DB 544 GAAAGGAGTGTGGCCAAAGAAAGACACACACCTTGGACAAATTTATATACCTTTTGTCT 603  
QY 659 gtaattgatgaaggaaagttgtgcactcagaacaaagaactccttgatgcaagatag 718  
|||||  
DB 604 GTATTGTATGAAGGAAAGTTGGCACTCAGAAACAAAGAACTCTTGATGTCAGGATAGG 663  
QY 719 gatgtgcattctgtcgggctgtgacctaaatgtcacacagtcacatgtgtatgtataacag 778  
|||||  
DB 664 GATGCTCATCTGCTCGGGCTGGCTTAAATGCACACAGCATTAAGTGTATGTAAACAGG 723  
QY 779 tctctg---ccaggtctgattgtgagtcacaaggaatcagttctatgtgcattgtatgga 835  
|||||  
DB 724 TCTGTGCTACCAAGGTCTGATTGTGATGTCACAGGAATCAGCTATTTGGCAGTGTATTGGA 783  
QY 836 atgggacacactcctgaagtgcaactcaatatctcctgaagtgcaacatctctgttag 895  
|||||  
DB 784 ATGGGACACACTCTCTGAAGTGCATCAATATTCCTCGAAGTCAACATTTCTTGTAGG 843  
QY 896 aaacatctgcgaagcgctctctgtgaaactcgcacaaataacttcttactgtcacaacatc 955  
|||||  
DB 844 AACCATGCGCGCGCTCTGTGAATCTCGCATTAATCTTCTTACTGTGTCAAACACTC 903  
QY 956 ttgttgaccttggaacggttttctactgttttgcataatcttctcccaacaatagatggc 1015  
|||||  
DB 904 TTGTATGACCTTGGACAGTTTCTACTGTGTTGTGATCTCTTCCACCAACAGATGATGC 963

QY 1016 atggaagcttatgtcaaaagttagacagctgtccagaaggaaccccaactacagatgaataat 1075  
|||||  
DB 964 ATGGAAGCTTATGTCCAAGTATGACAGCTGTCCAGAGAACCCCACTACGATTAATAAAT 1023  
QY 1076 aatgaagaagcgggaagacatgatgatcttactgattcttgaaatgtatgtgttcaag 1135  
|||||  
DB 1024 AATGAAGAAGCGGAAGACTATGATGATGATCTTACTGATCTGAATGATGTGTGTCAGG 1083  
QY 1136 ttgatgatgaacactctctctctcttataccaatttgcctcagttgtgccaagaagcatctc 1195  
|||||  
DB 1084 TTTGATATGACAACCTCTCTCTTATTCCAAAATTTGCTCAGTTGGCCAAGAACATCTCT 1143  
QY 1196 aaactctgggtacatacatgtcgtcgtgaagaagagctgtggaactgtcctctgtgc 1255  
|||||  
DB 1144 AAACTTGGGTACCTTATCTTGTCTGAAAGAGAGACTGTGGACTATGCTCTCTTATGTC 1203  
QY 1256 ctgcgcccgatgtgacagaagttataaagttcaatatttgaaacatgtgacctcagcgatt 1315  
|||||  
DB 1204 CTCGCCCCCGATGACAGAAAGTTATAAAGTCAATATTTGAACAATGGCCCTCAGCGATT 1263  
QY 1316 ggtaggaagtacaaaagaagtcgattatgtgcatacagaatgaacacttaagaactcgt 1375  
|||||  
DB 1264 GGTAGGAAGTACAAAAGTCCGATTTATGGCATACAGATGAACCTTTAAGACTGCT 1323  
QY 1376 gaagctatcagcatgaatcaggaatcttggaacttacttlatgtgggaagtgtgagac 1435  
|||||  
DB 1324 GAACCTTATCAGCATGTGATACGAAATCTTGGAACCTTACTTTATGGGAAGTTGGAGAC 1383  
QY 1436 aaactgttatataatttaagaatcaagaacagcagacataataactatcactcacaag 1495  
|||||  
DB 1384 ACACTGTTGATTATATTAAAGAAATCAAGCAAGCAGACCATTAACATCTACCTCCACGGA 1443  
QY 1496 atcacatgtatcgctcctcttggatccaagagatctacaaaagttgtaaaaacatttgaag 1555  
|||||  
DB 1444 ATCAGTATGTCCGCTCTTGTATTCAGGAGATTACCAAAAGGTGTAAACATTTTGAAG 1503  
QY 1556 gatttccaattctgcagcaggaataatccaatataatgtgaagctgactgtagaagat 1615  
|||||  
DB 1504 GATTTTCCAATTCGCCAGAGAAATATTCAATATTAATGACAGTGCATGTAAGAT 1563  
QY 1616 gggccaactaaatcaagatcctcgtgtcctgtgaccgctatactactagtttcgttaatatg 1675  
|||||  
DB 1564 GGGCCAACATAATCAGATTCCTCGGCTGACCGCATTAATCTAGTTTCTGTTAAATATG 1623  
QY 1676 gaaggaatctagtttaaggaatcattgtggccctctctactctgtcacaagaatctgta 1735  
|||||  
DB 1624 GAGAGATCTAGCTTCAAGGACTATTGGCCCTCTCTCATCTGTCTCAAAAGAAATCTGTA 1683  
QY 1736 gatcaagaagaaacagataatgttcagacaagaagaatgtcatcctgttttctgtatct 1795  
|||||  
DB 1684 GATCAAAAGAGGAACCAAGATATATGTACAGACAGAGATGTCACTGTTTCTGTATTTT 1743  
QY 1796 gatgaagaacgaagctgtgtaccttcacagagaatatacaagcttctccccaatccagct 1855  
|||||  
DB 1744 GATGAGAAACGGAAGCTGTGACTCAGAGAAATATACAGGCTTTTCCCAATCCAGACT 1803  
QY 1856 ggaatgagcttgaagatccaaggttccaagccccaactccaactgaagcaatcaatggc 1915  
|||||  
DB 1804 GGAAGTCAAGCTTGAGATTCAGAGATTCAGAGCTCCAAACATCAATGACACAGCATTAAGTG 1863  
QY 1916 tatgttttatagttgtcagttgtcagttgttttgatagagtggtgacatctgttatact 1975  
|||||  
DB 1864 TATGTTTGTATAGTTGTGCAAGTTGTGATTTGTGATGATGAGTGCATACTGGTATATT 1923  
QY 1976 ctaagcatgtgaagcaagactgacttcttctgtcttcttctctctgtatataacttcaaa 2035  
|||||  
DB 1924 CTAAGCATTTGGAGACAGACTGACTTCTTGTGCTCTTCTCTGATATACCTTCAAA 1983  
QY 2036 caaataatgtctatagaagcaacatcaacctatctccatctcagaagaacatgcttc 2095  
|||||  
DB 1984 CACAAATTTGCTATGTGAAGACACACTCACCTATTTCCCATTTTCAGGAGAACTGTCTTC 2043  
QY 2096 atgtcgaatggaataaccaggtctatagttatctgtggtgtgcacaactcagacttctggaac 2155



```

2044 ATGTCGATGGAAAAACCCAGGCTCTATGATTCGGGGGTGCCAACCTCAACACTTTGGGAAAC 2103
2156 agagagcatgacgcgccttactgaaggtttctgaqtlgtgaacaaagacgtygattatcac 2215
2104 AGAGGCAATGACCGCTTCTGTAAGGTTTCTAGTTGTGACACAGAACACTGCTATTATTAC 2163
2216 gagagacgtlta tgaagatattcagacatactgctgagtaaaaaaagatgccatlgaaaca 2275
2164 GAGGACAGTTATGAAAGATATTTCACGACTACTGTGAGTAAAAAACATGCCATTTGAAACA 2223
2276 agaagcttctcccgaaatcaagacaccctagcactagcgaaagaactttaatgcacc 2335
2224 AGA----- 2226
2336 ccacacagcttggaaacgcatacaacgggaataactcgtaactcttcagtcagatcaaa 2395
2227 -----GAAATTACTCGTACTACTTTCAGTCAGATCA 2259
2396 gaggaatctgactatgataccaatacagtlgaatgaagaagaagatttgacatt 2455
2260 GAGGAATTTGACTATGATGATACCATATCAGTTGAAATGAAGAAGATTGTCATT 2319
2456 tatgatgagatgaataatcagaagcccgccagcttcaaaagaaaaacgcacataattt 2515
2320 TATGATGAGATGAAAAACAGAGCCCGCACACTTTCAAAAACAGACACTATTATTT 2379
2516 atgtcagatggagaagctctggagatatgagtagtgcctcccaatlttcaaga 2575
2380 ATTCTGCAAGTGGAGAGGCTCTGGATTTATGGGATGAGTAGCTCCCCACATTTTCTAAGA 2439
2576 aacagagctcagagtgagcagtgctccctcagtlcagaagaagtgtttccagaatctac 2635
2440 AACAGGCTCAGAGTGGCAGTGTCCCTCAGTTCAAGAAAGTGTTCAGAGAAATTACT 2499
2636 gatgctcccttactcagcccttataccgttgagagaactaaatgaacatttggactcctg 2695
2500 GATGGCTCTTACTCAGACCCCTTATACCGTGAAGACTAAATGAACATTGGGACTCCG 2559
2696 gggacataataagagcagaagtgagaataataatcagtgtaacttcagaatacagcc 2755
2560 GGGCATATATTAAGACGCAAGAGTTGAATATATCATGTGTAACTTTCAGAATTCAGGCC 2619
2756 tctgcctcctatctcccttattctagccttatttctatlgaggaagatcagaagcaaga 2815
2620 TCTGTGCTCTTCTTCCTTATTTCTAGCCCTTATTTCTATGAGAGATCAGAGCAAGGA 2679
2816 gcaagaactagaaaaaaacttgtcgaagcttaatgaaccaaacttacttctggaaagt 2875
2680 GCAGAACCTTGAAAAAAATTGTCAAGCTTATGTAAGAACCAAAAACTTACTTTGGAAAGTG 2739
2876 caaatcatatcagcaccactaaagaatgatttgactgcaaaagccttgagcttattcct 2935
2740 CAACATCATATAGGCAACCCACATTAAGATGAGTTTGACTGCAAAAGCCTTGCGCTTATTTCT 2799
2936 gatgttgacttgaaaaaagatgtagcactcagcctgatttgagccctctcgtgctgcagc 2995
2800 GATGTTGACCTGGAAAAAGATAGTGCATCGACGCTGATTTGAGCCCTTGTGTCGCCAC 2859
2996 acttaaacacatgaaacccctgctcattgaggaacaagtgcacagtcagagaatttgcctgttt 3055
2860 ACTTAACACACGAAACCCGTCATGGGAGACAAAGTACAGTACAGAGATTTCCTGTGTTT 2919
3056 t----tcacacatcttggatgagaccaaagcttggtacttcaactggaataatatlgaaagaac 3112
2920 TTCTCTACACATCTTTGATGAGACCAAAAGCTGGTACTTCTACTGAAATAATGGAAGAAAC 2979
3113 tgcagaggtcccttgaaatatacagaatggaagaatccacattttaaagagaatatgccttc 3172
2980 TGCAGGGCTCCCTCGAATATCCAGATGGAAGATCCCACTTTTAAAGAGAAATTATGCTTC 3039
3173 catgcaatcaatggtctacataatgatacactggtcctgaagtaatggtcagagataaa 3232
|||||
3040 CATGCAATTCATGGCTACATATAATGGATACACTACCTGGCTTAGTAATGGCTCAGAGTCAA 3099
3233 aggattcgaatggtatcgtctcagcaatgggcagcaatgaanaacatccattcatcttc 3292
3100 AGGATTTCAGTGTATCTGCTCGCATGTGGCAGCATGAAACATCATTCTATTTCATTTC 3159
3293 agtgcacatglttcaactgtacgaaaaaagagagatlaaaatgagcctgtatacatctc 3352
3160 AGTGACATGTGTCTACGTACGAAAAAAGAGGAGTATTAATAATGGCACTGTACAATCTC 3219
3353 tatcagagtglttttgagaacagtggaataatgtaccatccaaagctgaaatttgcggtg 3412
3220 TATCCAGGCTTTTGAACAGTGAATGTAATGTACCATCCAAAGCTGGAATTTGGCGGGTG 3279
3413 gaatgccttatgaggagacatcatalcagcttggaatgaacaacttlttcgtgtacagc 3472
3280 GAATGCCCTTATTTGGGAGCATCTTACATGCTGGGATGAGACACACTTTTCTGTGTACAGC 3339
3473 aataagtgatcagactcccttgggaatggtctctggaacatalgaagatlttcagatlaa 3532
3340 AATAAGTGTCAAGACTCCCTGGGATGGCTTCTGGACACTTAGAGATTTCAGATTACA 3399
3533 gcttcagaacaatatgagcagtggtggcccaaaagctggtgcagacttcatatccgagatca 3592
3400 GCTTCAGACAAATATGACAGTGGGCCCAAAAGCTGGCCAACTTATTCGGATCA 3459
3593 atcaatgccttgagacacaagaagcccttctctggaatcagaagtgatctgttgacaca 3652
3460 ATCAATGCTGTGAGACCAAGAGAGCCCTTTCTTGATCAAGTGTGATCTTTGGACCA 3519
3653 atgattatcagcagcatcaagaacccaaggtgcccgtcagaagttctccagcctctacatc 3712
3520 ATGATTATTTCAAGCATACAGCATACAGCCAGGCTGCCGTCAAGAGTTCTCCAGCCTCTACATC 3579
3713 tctcagttatcatatgatatgattgttgtatgaggaagtgagagacttaccgaggaact 3772
3580 TCTCAGTTATCATATGATATGATGTCTTATGGAAGAGTGGACACTTATTCAGGAAAT 3639
3773 tccacttgaacacttaatgtcttcttcttgcaaatgtagatcatcttggaataaacaacat 3832
3640 TCCACTGGACCTTATATGATGTCTTTTGGCAATGTGATGATTCATCTGGGATTAACACAT 3699
3833 attttaaaccctcgaattatgtctgcatacatccglttgcaaccaactcatatagcat 3892
3700 ATTTTAAACCTCCCATATTATGCTCGATACATCCGTTTGGCAACCACTAATTATAGCATTT 3759
3893 cgcagcactcttcgcatgagatgtagtggatggatgagtttaaaatagttcagaatccatg 3952
3760 CGCAGCACTCTTGCATGGAAGTGAATGGGCTGTGATTTAAATAGTTGACGATGCCATTG 3819
3953 ggaatgagagtaaaagcaatalcaagatgcagagatctacgtctcatcctacttaccat 4012
3820 GGAATGGAGAGTAAGCAATATACATACATGACAGATTAAGTCTCATCTCATCTTATTCACAT 3879
4013 atglttgcacactggtctccttcaaaagcttcgaacttcaactccaaagggagagtaatgcc 4072
3880 ATGTTTGCACACTGCTGCTCTTCAAAAAGCTCGACTTCACCTCCAAGGAGGAGTAAGGCC 3939
4073 tggagacctcagtgbaataatccaaaagagtgctgcgaagtgagaggtctccgaagaacatg 4132
3940 TGGAGACCTCAGGTGAATAATTCAAAAGAGTGGCTGCAAGTGGACTTCCAAAGACATAG 3999
4133 aaagtccagagagtaactcaacgaggaatlaaaactcctgcgttaccagcaatgatatgaag 4192
4000 AAAGTCACAGAGATTAAGTACAGGAGTAAATATCTCGCTTACAGACATGTATGTGAAG 4059
4193 gagttccctcatctccagcagttcaagaatggtccatcagtgagactccttlttcagaatggtc 4252
4060 GAGTTCCTCATCTCCAGACATCAAGATGCGCATCAGTGAAGTCTCTTTTCAAGAAAGGCC 4119
4253 aaagtaaaaggtlttlttcagggaaatcaagaactccttcaacactgtggtggaactccttagac 4312
4120 AAAGTAAAGGTTTTCAGGAAATCAAGACTCTTCAACACTGTGGTGAACCTCTAGAC 4179
|||||
```



OY	4313	ccaccgttactgactcgcgtactccttgcgaaltcaccccccagattgggtgcaccagattgcc	4372
Db	4180	CCACCGTTACTGACACTCGCTACTCTTGAAATTACACCCCCAGAGTTGGGTGCACCCAGATTGCC	4239
OY	4373	ctgaagatgagaggttctctgtgctgcgagagccacagagaccctac	4414
Db	4240	CTGAGAGTAGAGGTTCTGTGGCTGCGAGGCACACGACCTCTAC	4281
RESULT	5		
LOCUS	108642		
DEFINITION	108642	4548 bp	PAT
ACCESSION	108642	Sequence 2 from Patent WO 8800831.	02-DEC-1994
VERSION	108642.1	GI:588648	
KEYWORDS			
SOURCE			
ORGANISM		Unknown.	
REFERENCE		Unknown.	
AUTHORS		Unclassified.	
TITLE		1 (bases 1 to 4548)	
JOURNAL		Pasek M.P.	
FEATURES		DNA SEQUENCES CODING FOR MODIFIED FACTOR VIII:C AND MODIFIED FACTOR VIII:C-LIKE POLYPEPTIDES AND PROCESSES FOR PRODUCING THESE POLYPEPTIDES IN HIGH YIELDS	
SOURCE		Patent: WO 8800831-A 2 11-FEB-1988;	
		Location/Qualifiers	
		1..4548	
		/organism="Unknown"	
BASE COUNT	1344 a	1004 c	1002 g
ORIGIN		1197 t	1 others

Query Match	92.5%;	Score 4085;	DB 10;	Length 4548;
Best Local Similarity	95.1%;	Pred. No. 0;	Mismatches	
Matches 4325;	Conservative	0;	Indels 31;	Gaps 3;
Oy	59	gccaccagaagaatctactactggtggtgcagttggaacgtcgaatgaggaactatagcaagat	118	
Db	1	GCACCAAGAATACTACTCTGGTCCACTGGACACTGATGAGACTATATGAGCAAGATAT	60	
Oy	119	ctcggtgagctgctctgtggaacgaagattctctctcagtgagtcgcaaaaatcttctccattc	178	
Db	61	CTCGGTGAGCTGCCCTGTGGACGCAAGATTTCTCTCTAGAGTCCAAATCTTTCCATTC	120	
Oy	179	aacacctagtcgtgtgtaaaaaagactctgtttgttagaattcaaggttcacctttcaac	238	
Db	121	AACACCTCAGTCGGTGAACAAAAGACTCTGTTGTAGAAATTCAGGAGATCACCTTTTCAC	180	
Oy	239	atcgctaagcccaagccaccctctgatatggtctgtcgaagttccatcatccaaagctgaagtt	298	
Db	181	ATCCGTTAAGCCAAAGGCCACCTCGATGGGTCTGCTAGTCTCTACCATCCAGGCTGAGGTT	240	
Oy	299	tatgatacagtggttcattacacttaagagacatggcttcccatctcgttcaagcttcattcct	358	
Db	241	TATGATACAGTGGTCAATTACACTTAAGAACATGGCTTCCATCCTGTGAGCTTCATGCT	300	
Oy	359	gttggtgtatctctcaactcgtaaaagctctctgaaggagctgtaatatgatgacagaccagtcaa	418	
Db	301	GTTGGTGTATCTCTACTGTAAGACTTTCAGGAGACTGTAATATGATGATCAGACCAAGTCAA	360	
Oy	419	agggaagaagaagtatgataaagtcttccctctgtgtgaagccatataatgtcttgcgaagtc	478	
Db	361	AGGGGAAGAAGATGATAAAGTCTTCCCTGCTGGAAACCCATAATATGCTGGCAGGTC	420	
Oy	479	ctgaagaagaaatgtgccaatgtgctcttgacccaactgtgccttaactaatatcttcttct	538	
Db	421	CTGAAGAAGATGTGGCCAAATGGCCTGTACCCACTGTGCTCTTACTACTCATATCTTTCT	480	
Oy	539	catgtgagcctgtgtaaaaagacttgaattcaaggtcctcatcttgagccctactagtatgata	598	
Db	481	CATGTGAGCCTGTGTAAGAAGCTTAATTCAGAGCCTCATATGAGCCCTACCTGTATGTGGA	540	

OY	559	gaaggagctctgcgcgaaggaaagacaagaccttgcacaatttactactcttctgtc	658
Db	541	GAAGGAGCTCTGGCCCAAGGAAAGAAACACACACCTTGCCACAAATTTATCTACTTTTGGCT	600
OY	659	gtatttgaaggaaagaatttgcactcagaaacaaagaccttgccttgcgaatag	718
Db	601	GTATTGTGTAAGGGAAGAAAGTTGGCACTCGAAGAACAGAACTCCTTGATGCAAGGATMAG	660
OY	719	gatcgtcatctctgcggccctgcgcctaaatgcacacagtcacatgttcatgtlaaacag	778
Db	661	GATGCTGATCTGCTGGCGCCTGGCCCTAAATGACACACAGTCAATGTTATGTAACACAG	720
OY	779	tctctg---ccaggtctgatgtgatgcacaggaatacagtcatactgtgcagtgtatga	835
Db	721	TCTCTGCTACCAAGCTCTGATGTGATGCCACAGGAAATCATGCTATTGGCATGTGATTGA	780
OY	836	atgggaccactcctgaagfcaetcaabtlctctgaagfcaacacttctgttggag	895
Db	781	ATGGGCACACTCCTGGAAGTGCATCAATATTCCTGGAAGTGCACATTTCTTGAGG	840
OY	896	aacatgcgcagcgctccttggaaatctgcacaatacttcccttactctgcacaacac	955
Db	841	AACCATGCCCAAGGGGTCTTGGAATATGCCCAATTACTTTCTTACTCTCAACACCTC	900
OY	956	ltgabgaccttggacagtttctactcgttttgcatactccttcccaacaatgatgyc	1015
Db	901	TTGATGACCTTGGACAGTTTCTACTGTTTGTCATATCTCTTCCACCAACATGATGTC	960
OY	1016	atggaagcttatgtccaaagttagaacgtgtgccagggaaacccaactacgaatgaaat	1075
Db	961	ATGGAAGCTTATGTCAAAAGTAGACAGCTGCCAGAGGAACCCCACTACGAATGAAAT	1020
OY	1076	aatgaagaagcggaagacctatgatagatccttaccgattctcgaatgtgaatgttgcag	1135
Db	1021	AATGAGACAGCGGAAGACTATGATGATGATCTTACTGATTTGAAATGCAATGGATGGTCAG	1080
OY	1136	ttgatgatgaacaactctccttcccttalcacaatctgcgtcagtlgtccaaagaatcct	1195
Db	1081	TTTGATGATGACACACTCTCCCTTCCTTATTCCAATTTGCGTCAAGTTCGAAGACATCT	1140
OY	1196	aaaacttgggttaacttatacttgcctgcctggaagggagactctggagctatgtctccctatgc	1255
Db	1141	AAAACCTGGGTACCTTTCATTTGCTGCGTGAAGAGAGACTGGGACATATCTCCTTAATGC	1200
OY	1256	ctgcgccccggtgacagaagtltaaagaacatcttgaacaatggccctcaagcgatc	1315
Db	1201	CTCGCCCCCGATGACACAAGTTATAAAGTCATATTGTAACATATGCCCTCAGCGGAT	1260
OY	1316	gttgaagaagtacaaaagtcgcgatttatgycatacacaatgtaaaccttlaagactcgt	1375
Db	1261	GGTGAGGAAGTACAAAAGATGCCGATTATGTCATACACAGATGAACCTTTAAGACTCGT	1320
OY	1376	gaagcattatcagcatgaaatcaagaaacttgggaacttactttaggggaagtggagac	1435
Db	1321	GAACCTATTTCAGCTATGTAATGAGAAATCTTGCGACCTTTACTTTATGGGGAAGTGGAGAC	1380
OY	1436	acaacttgtatataatttaaagaatcaagaagaagccataaacttcaaccttaagga	1495
Db	1381	ACACTGTGTATTATTTAAGTAATCAAGCAAGAGACCATATTAACATTCACCTTACGGA	1440
OY	1496	atcacgtatgtcgccttctgttatacaaggagatlaaccaaaaggtglaaaacattgag	1555
Db	1441	ATCACTGATGTCCCTCTTGTGTATCAAGSAGATTACAAAAGTGCTAATAACATTTGAAG	1500
OY	1556	gattttccaaatctgcacagagaataatccaataaataatgacagtgatctgtagaagat	1615
Db	1501	GATTTTCCAATTCTGCACAGAGAAATATTCAAATATAAATGACAGTGCATGAAGAT	1560
OY	1616	ggcgcaactaaatcaagctcctctgggtgcctgcagccgctatactctcagtttgcgttaatg	1675
Db	1561	GGGCGCAACTAATATAGATCTCGGTGGTGCACCCCGCTATTAATCTTAGTTCGTATATATG	1620
OY	1676	gagcgaatctcagcttcaagaccatcttggccctctccatctcgtctacaaaatacttata	1735



```

Db 1621 GAGAGAGTCTAGCTTCAGGACTCATGTGCCCCCTCTCATCTGCTACAAAGAACTCTGTA 1680
Qy 1736 gatacaagaagaaacacgaataatgtcagacaagaagaatgtcatctgtttcttatt 1795
Db 1681 GATTAAGAGGAACCAAGATATGTCAACAAAGAGAAATGTATCTCTGTTTCTGTATT 1740
Qy 1796 gatagaacccaagcttggtaacctcacaagaataatacaagcttctccccaatccagct 1855
Db 1741 GATGAGAACCAGAGCTGTACTCTACAGAAATATACAACGCTTCTCCCAATCCAGCT 1800
Qy 1856 ggaatgcagcttgaagatccagagttccaagcctccaatacaatgaagaatgaatgac 1915
Db 1801 GGAATGACGCTTGAAGGATCCAGAGTTCCAAAGCCTCAACATCATCAGCAGCATCAATGC 1860
Qy 1916 taatctttgtaagtgtcagttgtcagttgtgttgcataagtgaggatcagatcatt 1975
Db 1861 TATGTTTTGTAGTTTGTCAGTTGTGTCAGTTGTTGTCATGAGGTGGCATACTGTAAT 1920
Qy 1976 ctgaagcatlgaagacagactgaactctctctgtctctctctctcgtatataccttcaa 2035
Db 1921 CTAGCATTTGAGACACAGACTGACTTCTTCTGCTTCTCTCTGATATACCTTCAAA 1980
Qy 2036 cacaataatgtctatgaagaacactaaccttaattccattctcaggagaatcgtcttc 2095
Db 1981 CACAAATGGTCTATGAGAGACACACTCACCCCTATTCCATTCCTGAGAGAAACCTGCTTC 2040
Qy 2096 atgcgatggaacccccagctctatggaattctcgggtgcacaaactcagacttcggaac 2155
Db 2041 ATGTCGATGGAAGAACCCAGGCTCTATGATTTGCGGGTGCCACACTGACACTTTGGAAAC 2100
Qy 2156 agaggaatgaccgccttaactgaaggtttctagttgtgaacaagaacatggtgatactac 2215
Db 2101 AGAGGACATGACCGCCCTTCTAAAGGTTTCTAGTTGTGACAAAGAACACTGCTGATTATT 2160
Qy 2216 gaggaacgttataagatatatttcagcatactgtcgtgaataaacaatggccattgaaca 2275
Db 2161 GAGGACAGTTATGAGATATTTTACAGCATCTTGTGAGTAAAGAACATGCCATTGAACCA 2220
Qy 2276 agaagcttctcccaagat- 2293
Db 2221 AGAAGCTTCTCCAGGATCTCTTGTGCTGGATTAACCACTATGATACCAAAA 2280
Qy 2294 - 2293
Db 2281 GAAGATGGAATCCCAAGAGAACTCACAGAAAAACAGCTTTAAGAAAAAGATACC 2340
Qy 2294 - 2293
Db 2341 ATTTTGTCCCTGAACGCTTGTGAAGCAATCATGCAATAGCAGCAATTAAGAGGACAA 2400
Qy 2294 - lcaagacaccccttagcatctaggcaaaagcaatttaatt 2329
Db 2401 AATTAAGCCCAATATGAAGTCACTGCGGCAAAAGCAAGTAGACATGAAAGCTGTGCTCT 2460
Qy 2330 gccaccccaacagctctgaaagcccaatcaacgggaaataactcgtactactctcaatga 2389
Db 2461 CAAATCCACACAGCTCTTGAAGCGCCATCAACGGGAAATTAACCTGTACTACTTTCAGTCA 2520
Qy 2390 gatcaagaagaatctgactatgatatcaatcatcagttgaaatgaagaagaatctt 2449
Db 2521 GATTAAGAGGAATTTGACTATGATGATACCATATCATAGTTGAATATAGAGAGAAATTTT 2580
Qy 2450 gacattatgataaggaatgaataatcaagaccccgcaacttcaaagaanaacagacagac 2509
Db 2581 GACTTTATATGATGAGATGAAGAAATCAGAGCCCGCAGCTTTCAAAAGAAAAACAGACAC 2640
Qy 2510 taatttatgtctgagtgaggagagctctgggataatgggataagtagactccccacatgt 2569
Db 2641 TATTTTATGCTGCAGTGAGAGGCTCTGGGATTTATGGATGATGAGCTCCCCACATGTT 2700
Qy 2570 cttaagaacagggtctcagagtgaggagtgctccctcagttcaagaagctggtttccaggaa 2629

```

---

```

Db 2701 CTAAAGAACAGGGCTCAGAGTGGAGTGTCCCTAGTTCAGAAAGTGTTCCTCAGGAA 2760
Qy 2630 tttaactgagctcccttactcagccctataccgtggagaactaaatgaactattggga 2689
Db 2761 TTTACTGATGGCTCCTTACTAGCCCTTATACCGTGGAGAACTAATGAACATTTTGGGA 2820
Qy 2690 ctcctggggccataataagagcagaagtgaagataatcatatgttaacttcaagaat 2749
Db 2821 CTCCTGGGGCCATATATTAAGAGCAGAAAGTGAAGATATATCATAGTGAACCTTTCAGAAAT 2880
Qy 2750 caagcctctcgccctatttccttctatctagccctattcttcttgaggagaatcagaag 2809
Db 2881 CAGGCCCTCGTCCTATTCTTCTATTCTTATTCCTTATTTCTTATGAGGAAGATCAAGG 2940
Qy 2810 caaggaagacgaaccagaaaaaacttgcagagccttaatagaacccaacttactttgg 2869
Db 2941 CAAGGAGCAGAACTTGAAGAAACCTTTGTCAAGCCTATATGAACCAAACTTACTTTTGG 3000
Qy 2870 aaagtgcacaatcalatagaccacataaagatgaagttgaactgcgaagccctgggttat 2929
Db 3001 AAAGTCAACATCATATGACACCCACCTAAAGATGAGTTGACTCCAAAGCCTGGGCTTAT 3060
Qy 2930 tctctgattgtgaccttgaaaaaagtgtgcaactcagggcctgattgaaacctcttgct 2989
Db 3061 TTCTCTGATGTTGACCTGGAAAAAGATGTGACTCAGGCTGATTTGGACCCCTTCTGTGTC 3120
Qy 2990 tgcacacttaacacactgaacctgtctcatlgaaggagacagtgaaagtlacaggaattgct 3049
Db 3121 TGCCACACTAATACACTGAACTGCTCATATGGAGACAAATGAGATACAGAAATTTGCT 3180
Qy 3050 ctggtttt---tcaacatcttgaagaaccaaagctggttaacttcaactgaataatgaa 3106
Db 3181 CTGTTTTCCTCACCATTTGATGATGAGACCAAAAGCTGTACTTCTACAGAAATATGAA 3240
Qy 3107 agaaactcaggggtcccttgcataatccagaatgaagaatcccaacttttaagaagaattat 3166
Db 3241 AGAAACTGAGGGCTCCCTGCAATATCCAGATGGAGATCCCACTTTTAAAGAAATATAT 3300
Qy 3167 cgcctcatalgcataatagctgtactataatgaatacactcagcttgaatagtcacag 3226
Db 3301 CGCTTCATGCAATCAATGAGGTGACTAATATGATGATACACTGCTGATGATGAGCTCAG 3360
Qy 3227 gatcaagaatctgaatgtatctgtcgaagcatgggcagcaatgaataatcattctatc 3286
Db 3361 GATCAAGGATTCGATGATCTGCTCAGCATGGGACCAATGAAGAAATCATCATTTATT 3420
Qy 3287 caattcagtgacatgtgttcaactgtacgaaaaaagaaggagatlaaaatggcaacttac 3346
Db 3421 CATTTGAGTGGACATGTGTCTACTGTACGAAAAAAGAGAGTATTAATATGGCACTGTAC 3480
Qy 3347 aatcctataccagtgcttttgaagacagtggaatgttaccatccaaagctggaattgg 3406
Db 3481 AATCTCTATCCAGGTGTTTGTAGACAGTGAAGATGTATCATCATCAAAAGCTGGAATTTGG 3540
Qy 3407 cgggtggaatgccttaattggcgagacattacatgcttgggataagacaacttttctgtgt 3466
Db 3541 CGGGTGAATGCTTATTTGGGAGCATCTTACATCGCTGGGATGACACACTTTTCTGCTG 3600
Qy 3467 taacagcaatagtgacagactcccttgggaatggctcttgacacattagaagatttcag 3526
Db 3601 TACACATTAAGTGTCAACACTCCCTGGGAATGGCTTGTGACACATTAAGATTTTCAG 3660
Qy 3527 altacagcttcaagacaaatalggaagtgggccccaagcttgcagactcattatcc 3586
Db 3661 ATTACAGCTTTCAGCAATATGAGACAGTGGGCCCAAAAGCTGGCCAGACTTCATTATTCC 3720
Qy 3587 ggaatcaatcaatgtcctggaagcaaccaagagccctttcttgatccaaggtygactcgtgt 3646
Db 3721 GGATCAATCAATGCTCTGAGACACCAAGAGGCCCTTTCTTGATCAACAGGTGATCTGTTG 3780
Qy 3647 gcaaccaatgattatcaagacatcaagacccaagggtggcggtaaggaagtttccaagctc 3706
Db 3781 GCACCAATGATTATTCAGGCAATCAAGCCCAAGGCTGCCCTCAGAAAGTTCTCCAGGCTC 3840

```



3707 tacatctcctagttatcatcatatgatagtcttgatggaaagatggcagactatcga 3766  
|||||  
3841 TACATCTCTCAGTTATCATCATATATAGTCTTATGGAAGAAAGTGGCAGACTTATCGA 3900  
|||||  
3767 ggaattccactggaaacttaattgctctctcttgcaatgltgatactcctggataaaa 3826  
|||||  
3901 GGAATTTCCACTGGAACTTAATGCTCTCTTGGCAATGGATTCATCTGGGATFAAA 3960  
|||||  
3827 cacaatttttaacctccaataatgtctcgataacatcggttgcaaccaatcattac 3886  
|||||  
3961 CACAAATTTTTTAACCTCCCAATTTATGCTGATACATCCGTTTGCACCAACCATTAAT 4020  
|||||  
3887 agaatcgcaagactctctcgatggatgatggctgtgatttaataatgataagatg 3946  
|||||  
4021 AGCATTTGCCAGCACTCTTCGATGGAATGATGAGGCTGTGATTTAAATAGTTGAGCATG 4080  
|||||  
3947 ccattggaaatggaagtaaaagcaataatcagatgcacagatactgctctcatctt 4006  
|||||  
4081 CCATTTGGGAATGGAGATTAAGAATATCAGATGCACAGATTAATGCTTCACTCTTACTTT 4140  
|||||  
4007 accaatatgttgcacactgtctcctctcaaaagctgactccaactccaaggagagat 4066  
|||||  
4141 ACCAAATATGTTTGGCACCTGGTCTCTTCAAAAGCTGCACTTCACCTCAAGGAGGAGAT 4200  
|||||  
4067 aatggcctggagagactcaggtggaataatccaaagagtgctgcaagtgagactccagaag 4126  
|||||  
4201 AATTCCTGGAGACCTCGAGGAATTAATCAAAAGATGGCTGCAAGTGAGACTTCCAGAAAG 4260  
|||||  
4127 acaatgaagaatcacaaggaatgaactactcaaggagtaaatctctgcttaccagatgat 4186  
|||||  
4261 ACATGAATATCAACAGSAGTAATCACTCAAGSAGTAATAATCTGCTTACCAGATGATAT 4320  
|||||  
4187 gtaagaagttcctcatctcaccagatcaagaatgagcactaagtgagactccttcttccag 4246  
|||||  
4321 GTGAGAGATGTTCTCATCTCAGACAGTCAAGATGGCCATGAGTGAAGTCTCTTTTTCAG 4380  
|||||  
4247 aatggcaagaatgaaggtttctcagggaaatcaagaactccttcaactctgtggaactct 4306  
|||||  
4381 AATGGCAAAATGAAGGTTTTCAGAGGAATCAAGACTCTTCACTGAGTGAAGTCTCT 4440  
|||||  
4307 ctgaaccacacgttactgactgctactcctcgaatltcaaccacagatltggtgacacag 4366  
|||||  
4441 CTGACCCACCGTTACTGACTCGCTACTCTGCAATTCACCCCGAGATTGGTGACACAG 4500  
|||||  
4367 attgacctgagatgagatgtctggtgctcgagagcagacagactctac 4414  
|||||  
4501 ATTGCCCTGAGATGAGGTTCTGGCTCGAGGACACAGGACCTCTAC 4548  
|||||  
RESULT 6  
LOCUS 108641 4551 bp PAT 02-DEC-1994  
DEFINITION Sequence 1 from Patent WO 8800831.  
ACCESSION 108641  
VERSION 108641.1 GI:588647  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 4551)  
AUTHORS Pasek, M. P.  
TITLE DNA SEQUENCES CODING FOR MODIFIED FACTOR VIII:C AND MODIFIED FACTOR  
POLYPEPTIDES IN HIGH YIELDS  
Patent: WO 8800831-A 1 11-FEB-1988;  
JOURNAL location/Qualifiers  
FEATRES 1. 4551  
BASE COUNT 1345 a 1003 c 1004 g 1198 t 1 others  
ORIGIN

Query Match 92.5%: Score 4085; DB 10; Length 4551;  
Best Local Similarity 95.1%: Pred. No. 0;  
Matches 4325; Conservative 0; Mismatches 31; Indels 192; Gaps 3;  
QY 59 gccacagaagatactactccttggtgacgttggaactgtcatggaactatgcaaatgagat 118  
|||||  
Db 4 GCCACAGAAAGATATCACTGCTGGTGACAGTGAAGTGCATGAGGACTATATGCAAAAGAT 63  
|||||  
QY 119 ctgggtgagctgctctgtggaacgaagatctcctcctcgaagtgtccaaatcttccattc 178  
|||||  
Db 64 CTGGGTGAGCTGCTCTGTGGAGCGCAAGATTTCTCTGAGAGTCCAAAATCTTTTTCATTC 123  
|||||  
QY 179 aacacccagctcgtgtacaaagaactcgttgtgaaattcaagcttcaacttcaac 238  
|||||  
Db 124 AACACCTCAGTCTGTCAAAAAGACTCTGTTTATGAATTCACGATACACCTTTTCAAC 183  
|||||  
QY 239 atcgctaaagccaaagccacacttgatgtgctgctgaagctccatcacatccagcgaagtt 298  
|||||  
Db 184 ATGCTTAAGCCAAAGCCACCTGGATGGTCTGCTAGTCTTACCATCCAGGCTGAGGTT 243  
|||||  
QY 299 tatgatacagtggtcaatlaactaagaacatggtctccactcgtcagltcattgct 358  
|||||  
Db 244 TATGATACAGTGGTCATTTACACTTAAGAACATGAGCTTCCATCTGTCAGTCTTCATGCT 303  
|||||  
QY 359 gttagtgatcctacttgaaagctctcgaaggagctgaatatgtaagacacagca 418  
|||||  
Db 304 GTTGTGTATCTTACTGTGAAGCTTCTGAGGAGCTGAATGTGATGATCAAGACCACTAA 363  
|||||  
QY 419 agggagaagaagaatgaataagcttcccttggttgaaagccataatgltctggaagct 478  
|||||  
Db 364 AGGAGAAAGAAAGATGATTAAGTCTTCCCTGTTGGAAGCCATATATGTCGCAAGTGC 423  
|||||  
QY 479 ctgaagaagaatggtccaatggtccttgaccactgcttactcactatattcttct 538  
|||||  
Db 424 CTGAAGAAGATGCTCAAGTGGCTCTGACCCACTGTGCTTACTATATCTTCT 483  
|||||  
QY 539 catgtgacctgtgaagaagcttgaaatcagagcctcattggaagccactcagatgtaga 598  
|||||  
Db 484 CATGTGACCTGCTGAAGAAGCTTGAATTCAGGCTCTATTGGAGCCCTACTAGTATGTA 543  
|||||  
QY 599 gaaggagctgtgccaagaagaacacagacacttgacaaattatactacttctgt 658  
|||||  
Db 544 GAAGGAGTCTGGCCCAAGAAAGACACACACTTGCACAATTTATTAATCTTTTGGCT 603  
|||||  
QY 659 gttattgataagggaaagtgtgcaactcagaagaacaaagactccttgatgagagataag 718  
|||||  
Db 604 GTATTGTATGAAGGAAAGTGTGGCTCAAGAAACAAAGAACTCTTGATGACAGATAGG 663  
|||||  
QY 719 gatgtgcatctgctggtggtcctgaaatgaaatgaaacagltcaatggttataagacag 778  
|||||  
Db 664 GATGCTCAATCTGCTCGGCTCTGGCTTAATAATGCACACAGTCAATGTTATGTAACAGG 723  
|||||  
QY 779 tctctg---ccaggtctgattgtagtgcacaggaatcaatgltatltgcatgltatgta 835  
|||||  
Db 724 TCTCTGTACAGAGTGTGATTTGATGATGCCACAGGAATCAATGATTTGATGATTTGGA 783  
|||||  
QY 836 atgggacacactccttgatgacactcaatatctcctcgaaggtcaacacattctgtgag 895  
|||||  
Db 784 ATGGGACCACTCTCTGAAGTGCACCTCAATATCTCTCAAGGTCCACACATTTCTTGAGG 843  
|||||  
QY 896 aacatcgcaagcgtctccttggaatcctgcacataactccttacttactgttcaaacctc 955  
|||||  
Db 844 AACCATGCGCAGGCGTCTTGGAATTTCCCAATTAATCTTCTTACTGCTCAAAACCTC 903  
|||||  
QY 956 ttgatgaccttgagacgttctactglttgttgcatactccttccccaacatgtagtgc 1015  
|||||  
Db 904 TTGATGACCTTGAGACGTTTACTAGTGTGTCATATCTCTTCCCAACCAATGATGAGC 963  
|||||  
QY 1016 atggaagcttatgtcaaatgataagacgtgtccagaggaacccaactaagatgaataat 1075  
|||||  
Db 964 ATGGAAGCTTATGTCAAAAGTGAACAGCTGTCCAGAGGAACCCCAACATTAAGTAAAT 1023  
|||||  
QY 1076 aatgaagaagcggaaagactatgatatgatacttactcgtatcttcaatgagtggtcag 1135



```
|||||
Db 1024 AATGAAGAGGGAAGATATGATGATGATCTACTGATTTGGAATGATGATGTCAGG 1083
QY 1136 ttgatgatgaacatccctccctcccttaaccaatctgcctgaatctgcaagaagatcct 1195
Db 1084 TTGATGATGCAACTCTCCCTCTTCTTATCCAAATTCCTCGTGGCCAGAGATCTCT 1143
QY 1196 aaacttggtacatacaatctgctgctgaagaagagactggactgctcccttagtc 1255
Db 1144 AAAACTGGGATACATTTACATTTGCTGCTGAGAGAGACTGGACTATGCTCCCTTAGTC 1203
QY 1256 ctgcgcccgatgacagaagtataaaagtcataatctgaacaatgcccctcagcgatc 1315
Db 1204 CTCGCCCCCGATGACAGAGATTATAAAGTCATATTGGAACAAAGGCCCTCAGCGGATT 1263
QY 1316 ggtggaagatcaaaaagctccgatttatgcatcacagatgaaactttaagaactcgt 1375
Db 1264 GGTAGGAAGTACAAAAGTCCGATTATGCGATACAGATGAACCTTTAGACTCGT 1323
QY 1376 gaagctatcagcaltgaatcagaatctctggacccttactctatctgagggaagttagagac 1435
Db 1324 GAAGCTATTGCGATGATGATGAGAACTCTTGGACCTTTACTTTATGGGAACTTGGAGAC 1383
QY 1436 acactgtgatataattagaatcaagcaagcagacataatacctacccctcagga 1495
Db 1384 ACAGTGTGATTATTTAAAGATCAACCAAGCAAGCATATTAACATCTACCCCTACGGA 1443
QY 1496 atccatgatgcgctccctcttctatcaaggagattaccaaaagtgctaaacatttgaag 1555
Db 1444 ATCACTATGTGCTCCCTCTTCTTATTCAGAGATTACCAAAAAGGTAAACATTTGAGG 1503
QY 1556 gatttccaatctcgcagagaaatatccaataataatgacagctgactgagaagat 1615
Db 1504 GATTTCGAATTCTGCCAGGAATAATTCAATTAATGAGACAGTGCCTGATAGAGAT 1563
QY 1616 gggcacaactaaatcaagatcctcggtgcctgacccgctattactcagttcgtaatatg 1675
Db 1564 GGGGCAACTAAATCAGATCCCTGGGCTCGACCCGCTATTACTTCACTTCTTAAATATG 1623
QY 1676 gagaagaatcagctctcagacatctgagcctcctccatctgctaaagaatctcta 1735
Db 1624 GAGGAAATCTAGCTTCAAGGACTCATTTGGCCCTCTCTCATCTGCTTCAAGAAATCTGTA 1683
QY 1736 gatacaagaagaaacagataatgtaagaagaagaaatgtaactcgtcttctgtaatt 1795
Db 1684 GATCAAGAGAGAAACAGATATGTCAGACAAGAGATGATCTCTGTTTCTGTATATT 1743
QY 1796 gatgagaacagaaactggttacctcaagaagaatacaagctcttctcccaatccagct 1855
Db 1744 GATGAGAACCCGAACCTGGTACTCTCACAGAAATATTAACAGCTTCTCCCAATGCAGCT 1803
QY 1856 ggaagtgaacttgaagatccagagatctccaagcctccaacatcatgacagacataatgac 1915
Db 1804 GGAATGAGCTTGAGATCCAGAGTTCAGAGCTTCCAAACATCATCATCAAGCATCAATGCG 1863
QY 1916 tatgttttgataagcttgcagcttgtaagcttgcttctcctcctcgtgatatccctcaaa 1975
Db 1864 TATGTTTGTGATGTTTGCACTTGTGCACTTGTGTTGATGAGGTGGCATTTAGCTATTAAT 1923
QY 1976 ctgaagcatggagacagaactgacttccctctcctcctcctcctcgtgatatccctcaaa 2035
Db 1924 CTAAAGCTTTGGAGCACAGACTTCTCTTCTGCTTCTCTCTGATATTAACCTTCAAA 1983
QY 2036 cacaanaatgtctctatgaagacacactccctatcccatctcagaagaactctcttc 2095
Db 1984 CACAAAATGGTCTATGAGACACACTCCCTATTTCCATTTCTCAGGAAACTGTCTTC 2043
QY 2096 atgtctgaatgaaaacccagctctataatgcttgcgggtgccaacaactcagacttggaaac 2155
Db 2044 ATGTGATGAGAAACCCAGGCTATGATTTCTGGGGTGCACAACTCAGACTTTTGGAAAC 2103
QY 2156 agaggaatgacgccttaactgaaaggttctcagcttgtaagaagaactggtgattattaac 2215
|||||

Db 2104 AGAGCATGACCGCTTACTGAGAGGTTTCTAGTTGTGACAAGAACACTGCTGATTTATAC 2163
QY 2216 gaggaacgtatgaagatatttcaagactactgctgagttaaaaacatgcatgaaaca 2275
Db 2164 GAGGACAGTTATGAAGATATTTTACACATACCTTGTGATTAACAAATGCTATTGAACA 2223
QY 2276 agaagctctcccaaat----- 2293
Db 2224 AGAAGCTTCTCCAGAGATCCTTCTGTTGGATACACACTATGTGACTCAGATACCAAAA 2283
QY 2294 ----- 2293
Db 2284 GAAGATGGAATTCACAGAGAGAGTCCACAGAAAAACAGCTTTTAAGAAAAAGATACC 2343
QY 2294 ----- 2293
Db 2344 ATTTGTCCTCGTAACGCTTGTGAAAGCAATCATGCAATAGCAGAAATTAAGAGGACMA 2403
QY 2294 -----tcaagacacccctagcactagcgaagaacttaaat 2329
Db 2404 AATAAGCCCGAAATPAGAAAGTCACTGGGCAAAAGCAGAGTAGACTGAAAGGCTGCTCT 2463
QY 2330 gccaccccaacagctctgtaaaagccatacaaggggaataaactcgttactccttcagta 2389
Db 2464 CAAAACCCACAGTCTTGAACGGCCATCAACGGGAAATTAACCTGTAATCTTCACTCACTCA 2523
QY 2390 gatcaagagaagaattgactatgatataccaatalcagtttgaatgaaagaagaattt 2449
Db 2524 GATCAAGAGGAATTTGACTATGATGATATCCATATCAGTTGTAATTAAGAAAGAAATTTT 2583
QY 2450 gacatttataatgaaatgaatcaagaccccgacgcttccaagaagaacacgaac 2509
Db 2584 GACATTTATGATGAGAGATGAAGAAATCAGAGCCCGACGCTTCAAAAGAAACACGACAC 2643
QY 2510 taattatgtcgcagtgagagagctcgtgattatagatgagatagctccccaatgct 2569
Db 2644 TATTTTATTTGCTGAGTGAAGGCTCTGGGATTTATGGAATGAGATAGCTCCCAATGTT 2703
QY 2570 ctaagaacaagagctcagagtggcagtgctccctcagttcaagaagaatgctttccagaa 2629
Db 2704 CTAAAGAACAGGGCTCAAGTGGCAGTCCCTCGAGTTCAAGAAAGTTGTTTCCAGAA 2763
QY 2630 ttactatgagctccttactcagccctataccgttggagaactaaatgaaacattggga 2689
Db 2764 TTTACTGATGGCTCCTTACTACAGCCCTTATACCGTGAGAACTTAATGAACATTTGGGA 2823
QY 2690 ctccctggggccatatataagagcagaagtgtgaagataatcatgtaacttgaagaat 2749
Db 2824 CTCTGGGGCCATATATTAAGACAGAAAGTTAAGATTAATCATATGATGTTATGAGAAAT 2883
QY 2750 caggcctctgctccctatccttctatctagccttattcttatatgagaagatcaag 2809
Db 2884 CAGGCTCTCGTCCCTATTATCTTCTATTTCTAGCCCTTATTTCTTATGAGAAATCAAGG 2943
QY 2810 caaggaagcagaacctagaaaaaacttgtaagccttaatgaacccaacttacttgg 2869
Db 2944 CAAGAGCAGAGAACCTPAGAAAAAATTTGTCAAGCTTAATGAACCAAAACTTACTTTGG 3003
QY 2870 aaagtcacaatcatatgagacccaactaaagaatgatttgctgcaagcctggactat 2929
Db 3004 AAAAGTCAACATCATATATGACCCCACTAAAGATGAGTTGACTGCTCAAAAGCCTGGCTTAT 3063
QY 2930 ttctctgattgtgaacttgaagaagaatgtaacccaagcctgacttggaccctctgctc 2989
Db 3064 TTCTCTGATGTTGACCTTGAAAAAATGTCATCAGGCTGATTTGGACCCCTTCTGCTC 3123
QY 2990 tgcacactaacacactgaacctgtctcatgaggagacaagtgtacagtlacagaaattgtct 3049
Db 3124 TGCCACACTTAACACACTGAACCTGCTCATGAGAGACAAGTGAAGTGAACAGAAATTTGCT 3183
QY 3050 ctgtttt---tcacacatcttgatgagaccaaagcctggtacttccactgaaaaatgtaa 3106
Db 3184 CTGTTTTCCTCACCATCTTGTGATGAGACCAAAAGCTGTGTTCTTCACTGAAAAATATGAA 3243
```



QY 3107 agaaactgcagggtccctcgaataatccagatggaagatcccaattttaaagaagaattat 3166  
|||||  
Db 3244 AGAAACTGCAAGGCTCCTCGCAATATCCAGATGSAAGATCCACTTTTAAAGAAATTAAT 3303  
QY 3167 cgtctcgaatcaatcaatggtctacataatgatactactcgtgcttagtaaatggtccag 3226  
|||||  
Db 3304 CGCTTCATGCAATCAATGGGTACATTAATGATACACTACTGGCTTAGTAATGATGATCAG 3363  
QY 3227 gatcaaaagattcgatggtatctgctcgaagatggcgaatgaaacatccattctattc 3286  
|||||  
Db 3364 GATCAAAAGGATTCGATGATCTCTCAGCATGGGCAATGAAACATCCATTCATTTAT 3423  
QY 3287 cattcagtgagacatggtctcactgtaacgaanaaagaagaatgataaagtcgacttac 3346  
|||||  
Db 3424 CATTTCAGTGACATGTTGTTCACTGTACGAAATAAGAGATTAATAATGGCACTGTAC 3483  
QY 3347 aatctctacaaagtgctttcttgagacagtggaatgttacatccaaagcttgaaattcg 3406  
|||||  
Db 3484 AATCTATATCCAGGTGTTTGTGAGACAGTGAATGTTACATCCAAAGCTGGAATTTGG 3543  
QY 3407 cgggtggaatgcttatttggcgaagatctacatgctggatgagacacttttctggtg 3466  
|||||  
Db 3544 CGGGTGAATGCCCTTATTTGGCGAGCATCTACATGCTGGATGAGCACACTTTTCTGTG 3603  
QY 3467 tacagcaataagtgctgaagctccccctgggaatggtctctgacacatgaagatttcag 3526  
|||||  
Db 3604 TACAGCAATTAAGTGTCACTCCCTGGGATGGCTTCTGACACATTAAGATTTTCAG 3663  
QY 3527 attacagcttcagacaataatggaacagtggtgccccaaagcttgccgaagacttcatatcc 3586  
|||||  
Db 3664 ATTAACAGCTTCAGAGCAATATGAGACATGGGCCCCAAAGCTGGCCAGACTTCATTAATCC 3723  
QY 3587 ggtatcaatcaatgctcgtggaagacaaagagaccccttcttgatcaaatggtggtatctg 3646  
|||||  
Db 3724 GGATCAATCAATGCTGGAGACACCAAGAGACCCCTTTCTTGATCAACAGTGGATCTGTG 3783  
QY 3647 gcaaccaatgatttaacacgagcatcaaaagacaaagtgccgctcaagaagcttctcagctc 3706  
|||||  
Db 3784 GCACCAATGATTAATTCACGGCATCAAGACCAAGGTCGCCGTCAAGATTTCTCCAGCTTC 3843  
QY 3707 tacatctcgaattatcatcatgatactgatacttgatggaagaatggcgaagacttaccga 3766  
|||||  
Db 3844 TACATCTCTCAGTTTATCATCATGTATAGTCTGATGGAAGAAAGTGGCAGACTTATCGA 3903  
QY 3767 ggaatctcactggaaccttaatggtctctcttggcaatggtgattcatctgggataaaa 3826  
|||||  
Db 3904 GGAATTCACACTGGAACTTAATGCTTCTTTGGCAATGTGATTCATCTGGCATTAATA 3963  
QY 3827 cacaatattttaacctcccaattatgtcgcgataacatccgittgaccccaaacatcatat 3886  
|||||  
Db 3964 CACAATTTTTTAACCTCCCAATTAATTTGCTCGATACATCCGTTTGCAACCACTCATTA 4023  
QY 3887 agcatctgcagcactcttcgcagatggaatgataggtcgtgatttaataatgttcgaatg 3946  
|||||  
Db 4024 AGCATTTGCGACACTTTTCCATGAGATGATGATGGGCTGTGATTTAATTAATGATGGACATG 4083  
QY 3947 ccaattgggaatggaatgaagaataatcagaatgacagattactgcttcaatccattc 4006  
|||||  
Db 4084 CCAATGGGAATGAGATTAAGCAATATCAGATGACACAGATTAATGCTTATCTTACTCTTT 4143  
QY 4007 accaataatgttgcacactgcttcccttcaaaagctggaattcaacctccaaaggaagaagt 4066  
|||||  
Db 4144 ACCAATATGTTTGGCACTGTCTCTCTTCAAAAGCTGCACTTCACTCCAAAGGAGAGT 4203  
QY 4067 aatgctcgaagacctcagtggaatcaatccaaagagtgctcgaagtggaacttccagaag 4126  
|||||  
Db 4204 AATGCTCGAGACCTCAGTGGAATTAATCCAAAGAGTGGCTGCAAGTGGAAGTCCAGAG 4263  
QY 4127 acaatggaatgcaaggaatgaataactcgaaggaatgaataatctctgcttaccagcatgat 4186  
|||||  
Db 4264 ACAATGAAGATCAAGAGATTAATCACTCAAGGAGTAATAAATCTGCTTACCAAGCATGTAT 4323

QY 4187 gtaagagatttcctcatctccagcagtcgaatggaatccatcagtgactcctcttccag 4246  
|||||  
Db 4324 GTGAAGAGATTTCATCTCCAGCAGTCAGATGCGCATCATGATGACTCTTTTTCAG 4383  
QY 4247 aatgcaaatgaaggttttcaaggaaatcaaacactccttcaacactgtgtgtaactct 4306  
|||||  
Db 4384 AATGCAAAAGTAAAGTTTTTCAGGAAATCAAGATCTCTTCAACACTGTGTGTAAGTCT 4443  
QY 4307 ctgaaccacacgcttactgactcgtactcctcgaaattcaaccccaagattgggtgacag 4366  
|||||  
Db 4444 CTAGACCCACCGTTTACGTACCTGCTACTTCTGAATTCACCCCGAAGTGGGTGCACAG 4503  
QY 4367 atgcccctgaagatggaaggtctggtcgtcgaggaacagaacctctac 4414  
|||||  
Db 4504 ATTGCCCTGAGAGATGAGAGTTCTGGCTGCGAGGACACAGACCTCTAC 4551

RESULT 7  
AR034084  
LOCUS AR034084 5035 bp DNA PAT 29-SEP-1999  
DEFINITION Sequence 1 from patent US 5869292.  
ACCESSION AR034084  
VERSION AR034084.1 GI:5949689  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 5035)  
AUTHORS Voorberg,J.,J.  
TITLE Hybrid proteins with modified activity  
JOURNAL Patent: US 5869292-A 1 09-FEB-1999;  
FEATURES  
source 1..5035  
BASE COUNT 1484 a 1127 c 1110 g 1314 t  
ORIGIN

Query Match 84.8%; Score 3743.8; DB 9; Length 5035;  
Best Local Similarity 87.5%; Pred. No. 0;  
Matches 4365; Conservative 0; Mismatches 52; Indels 570; Gaps 2;

QY 1 catgcaaatagagctcccaactgctctctctgctgcttggatctcgtcttaatgctc 60  
|||||  
Db 34 CATGGAATAGAGCTCTCCACTGCTTCTTGTGCTTGTGGATTCCTTTAGTGC 93  
QY 61 caccagaagatactactggtggtcagtggaactgcatggaactatgcaaatgactct 120  
|||||  
Db 94 CACCAGAAGATACCTGAGGTGAGTGAAGTGTATGAGACTATATGCAAAAGTGAATCT 153  
QY 121 cgtgtgagctgctgtggaacgaatcttccctcagaagtgccaaatcttccatcaa 180  
|||||  
Db 154 CGGTGACCTGCTGTGAGCGCAAGATTTCTCTAGAGTGCAGAAATCTTTTCCATTCAA 213  
QY 181 caactcagctggtgcaaaaagaactggtttagaattcaacgttcaaccttcaacat 240  
|||||  
Db 214 CACCTCAGTCTGTACAAAAAGACTCTGTTGTGAATTCAGGATACCTTTTCAACAT 273  
QY 241 cgttaagcacaagccacccctggaatggtctgtagtgcctacacatccaagctggaagttta 300  
|||||  
Db 274 CGCTTAAGCCAAAGGCCACCTGAGATGGGTGCTGAGTCTCAATCCAGGCTGAGGTTTA 333  
QY 301 tgatacagtggtcatcttaacttaagaacatggtctcccatccctgctgacttctatgctgt 360  
|||||  
Db 334 TGATACAGTGTCTATTAACCTTAAGAACATGAGTCTCCATCTCTGCTGCTGCTGT 393  
QY 361 tgtgtatctcactggaagctcttgaggagctggaatgatatgatatgaagccaatcaag 420  
|||||  
Db 394 TGGTGTATCTACTAGGAAGCTTGTGAGGAGCTGGAATATGATGATGACACCACTCAAG 453  
QY 421 ggagaaagaaatgataaagltctccctggctggaagcacaatataatgctcgaaggtcct 480  
|||||  
Db 454 GGAGAAAGAAATGATTAATGCTTCCCTGGTGAAGCCATATATGTCTGGAGGTCT 513



481 gaaagaaatggtccaatgacctgacacacatgaccttaacctatactatcttca 540  
514 GAAAGAAATGTCGAATGGGCTGTGACCACTGTGCTTACTACTACTATCTTTCTCA 573  
541 tctgacctggttaaaagacttgaatcagacctatctggaacctactagtctagaga 600  
574 TGTGGACCTGTGTAAAGACTGGAATTCAGGGCTCATTTGGAGCCCTACTAGTATGAGAA 633  
601 agggagcttgccaaagaaagacacagaccttgacaaatctatactactcttgcgt 660  
634 AGGAGTCTGGCCAGAGAAAGACACAGACCTTGACAAATTTATCTATCTTTGCTCT 693  
661 atttgatgaagggaaaatctggacctcagaacaaagaactctctgatgcaagatagga 720  
694 ATTGTAGTAAGGAAAGAACTTGGCACTAGAAACAAAGAACCTCTTGATGCAGATAGGA 753  
721 tctgcatctctgctggacctgacctaaatgcaacagtcacatggttatacaagctc 780  
754 TGTGCAATCTGCTGGGCGCTGGCCCTAAATGCACAGTCAATGGTTATGTAAACAGTCT 813  
781 tctgcaagctctgatttgatgcccacaggaatcagtcatacttgacatgcatggaatgg 840  
814 TCTGCCAGGTGTGATGATGCCACAGAAATCACTATTGGCATGTGATTTGGATGGG 873  
841 caccactctgaaatgacatcaatactctgaaaggtcacacattctctgtaggaacca 900  
874 CACCACCTCTGAAGTGACCTCAATATCTCTCGAAGTGACACATTTCTTTGTAGGACCA 933  
901 tctgcaagctctcttgaaatctcgcgaataacttctctactgctcaaacctctgat 960  
934 TCGGCAAGCGCTCTGGAAATCTCGCCAAATCACTTCTACTGCTCAACCTCTTGAT 993  
961 ggaaccttggaagaatttctactgcttctgcatatctctcccaacacatgtagcatgga 1020  
994 GGACCTTGAGACAGTTTCTACTGTTTGTCAATCTCTTCCACCAACATGATGGCATGGA 1053  
1021 agcttatgtcacaagtacagacgtctccagaggaaccccaactacgaatgaaataataga 1080  
1054 AGCTTATGTCAAAAGTAGACAGCTGTCCAGAGAACCCCACTAGCAATGAATAATATGA 1113  
1081 agaaagggaagaaactatgtagatctactgactcgaatggaatggtgtaggtttga 1140  
1114 AGAAGCGGAACATGATGATGATCTTGAATCTGAAATGATGATGCTGAGCTTTGA 1173  
1141 tgaatgaacactctctctcttataccaattcgcctcagttgccaagaagacatccaaac 1200  
1174 TGATGACAACTCTCTCTCTTATCCAAATTCGCTCAGTTGCCAAGAACATCTTAAAC 1233  
1201 tctggatcatatcatctgctgctgaaaggaggaactggaactatgctccctatgctcgc 1260  
1234 TTGGGTACATTAATGCTGCTGAAGAGAGAGACTGAGACTATGCTCCCTTATGCTCTGC 1293  
1261 ccccgatgaagaaagtataaaagtcaatatttgacaatgacctcgcggaattgtag 1320  
1294 CCCCAGTAGACGAAGATTATAAAGTCAATATTTGAACAATGGCCCTCAAGCGGATTGGTAG 1353  
1321 gaagtacaaaaaagctccgacttatgcaatacagaatgaaccttaagaactcgtgaagc 1380  
1354 GAAGTACAAAAAAGTCCGATTATGCGCATACAGATGAACCTTTAAGACTCGTGAAGC 1413  
1381 tattcagcatgaatcagaatctctgggaacttacttataggggaagtctgagacacact 1440  
1414 TATTCACATGAATCAGAAATCTTGGGACCTTACTTTATGGGGAAGTTGAGACACACT 1473  
1441 gttgatatatttaagaatcaagaacagacacataatacatctacccctcaagataac 1500  
1474 GTTGTATTATATTAAAGATCAAGCAAGACACATTAACATCTACCTCACGGAATCAC 1533  
1501 tgaatcgtctctcttgtaataaaggagataccaaaagggttaaaaaacttgaagatt 1560  
1534 TGATGTCCGCTCTTTGTATTCAAGGATTTACCAAAAGTGTAAGAACTTTGAAGGATTT 1593

1561 tcccaattctccaggagaataattcaaatataatlgacaagtgactgaagaatgggcc 1620  
1594 TCCAAATTCTGCCAGAGAAATATTCAAATATTAATGACAGTACTAGAAAGATGGGCC 1653  
1621 aactaaatcagatctctgggtgctctgaacccgctataactctagtcttataatgagag 1680  
1654 AACTAAATCAGATCTCGGTGGCTGACCCGCTATTACTGTAGTTCTGTTAATGAGAG 1713  
1681 agatcagctcagactcaatctggccctctccatctcgtcacaaagaactctgtagta 1740  
1714 AGATCTAGCTTCAGAGCATATGGGCTCTCTCTCATCTGCTACAAAGATCTGTAGATCA 1773  
1741 aagaggaaaccagataatgctcagaagaaggaatgcatcctgttcttgattatgta 1800  
1774 AAGAGAAACCGATATATGTGACACAAGAGAAATGTACTCTGTTTCTGTATTGATGA 1833  
1801 gaacggaagctgtagtaccctcaagaagaataataaagcttctccccaatccaagctgag 1860  
1834 GAACGGAAGCTGTGACTCTACAGAGAAATATCAACGCTTTCTCCCAATCCAGCTGAGT 1893  
1861 gcaacttgaggatccagaatccaaagctccacaatactgacaagatcaatggtctatgt 1920  
1894 GCAGCTTGAGATTCAGAGTTCCAGAGCTTCCAGCTCAACATCATGACAGCATATGGCTATGT 1953  
1921 ttctgatagttgcagttgctcagttgttgatgaaatggatggaactgtagatcaatlaag 1980  
1954 TTTGTATGTTTGGAGTTGTCTAGTTGTGATGATGAGTGAGCATATCTGATTTCAAG 2013  
1981 catggagcaagactgctctctctgtctctctctctctgatatactctcaaacaa 2040  
2014 CATTTGAGACACAGACTGACTCTTCTGCTCTCTCTCTGATATATACCTTCAACACAA 2073  
2041 aatgtctataagaacacaactcaacctatccactcctcaagaagaactgctcactgctc 2100  
2074 AATGCTTATGATGAAGACACACTACCCCTATTCCTCATCTCAGAGAAATGCTTCACTGC 2133  
2101 gatggaaaacccaagctctatgattctgagggtgccaactcgaactctcgaaacagag 2160  
2134 GATGAAAAACCCAGGCTATGATGATCTGGGGGCCACAACCTCAGACTTTCGGAACAGAG 2193  
2161 catgacgctctactgaaaggttctcagttgtgacaagaacactggtgattatctaggga 2220  
2194 CATGACCGCTTACTGGAAGGTTTCTAGTTGATTTCCAGAGGGGAGGAGCGACGACTA 2253  
2221 caatgaatgaatatca-----gcatactctgtagtaaaacaatgc 2265  
2254 TGTGACCTGTGAGATATTCAGTGAAGAGAGAGACTGACATCTGCGACACTCT 2313  
2266 catgaaccaagaagctctctccagaatctcaagacacctgacactagggcaagaactc 2325  
2314 GATTGAACCCAGAAAGCTTCTCCCAAAATTCAAAGACACCTTGACTAGGCAAAAGCAATT 2373  
2326 taatgcaac----- 2334  
2374 TATATCCACACCAATTCAGAAATGACATAGAGAAGACTACCTTGTTTGCACACAG 2433  
2335 ----- 2334  
2434 AACACCTATGCTAAATATCAAAATGTCTGCTAGTATTGTTGATGCTCTTGCGACA 2493  
2335 ----- 2334  
2494 GAGTCTACTCCACATGGGCTATCTTATCTGANTCTCCMAAGCAAAATATGAGACTTT 2553  
2335 ----- 2334  
2554 TTCTGATGATCATCACTGTGAGCAATAGACAGTAATTAACGCTGTCTGAATGACACA 2613  
2335 ----- 2334  
2614 CTTCAAGGCACAGCTTCATCAGAGTGGGACATGTTATTACCCCTGATGACGCTCCA 2673  
2335 ----- 2334



Db 2674 ATTAAGATTAAATGAGAACTGGGGACAACAGACAGATCCTCTGCTGGATTAACACTA 2733  
QY 2335 ----- 2334  
Db 2734 TGGTACTCAGATACCAAAAAGAGAGTGGAAATCCCAAGAGAGTCCACGAAAAACAGC 2793  
QY 2335 ----- 2334  
Db 2794 TTTTAAAGAAAAGATACCATTTTGTCCCTGACGCTTGTAAGCAATCATGCAATAGC 2853  
QY 2335 ----- 2334  
Db 2854 AGCAATAATATGAGGAGACAAAATAAGCCGAAATAGAAATGACCTGGGCAAGCAAGTAG 2913  
QY 2335 ----- 2370  
Db 2914 GACTGAAGGCTGCTCTCAAAAACCCAGCTTGAAAGCCATCAACGGGAAATATAC 2973  
QY 2371 TCGTACTCCTTCAGTCAAGTCAAGAGAGAAATGACTATATGATACCATATCATTTGA 2430  
Db 2974 TCGTACTCCTTTCAAGTCAAGAGAAATGACTATATGATACCATATCATTTGA 3033  
QY 2431 aatgaagaagaagatttgacattatgagatgaaatcaagagcccccagctt 2490  
Db 3034 AATGAAGAAGAGATTTTGACATTTATGATGAGATGAATAATCAGAGCCCCGAGCTT 3093  
QY 2491 tcaaaagaagaacagacacttttatgtcgcagtgagagagctcgggaattagat 2550  
Db 3094 TCAAAAAGAAAACAGACACTTTTATTTATGTGACAGTGAAGGCTCTGGGATTTAGGAT 3153  
QY 2551 gagtagctcccccacatgttctaaagaagaaggtccagatgagcagtgctccctacgttcaa 2610  
Db 3154 GAGTAGCTCCCCACATGTTCTTAAGAAACAGGGCTCAGAGTGGCTGCTTCAGTTCAA 3213  
QY 2611 gaaagttgttccagaagaatttactgagtgctccttactacgccttataccgltgaga 2670  
Db 3214 GAAAGTTGTTTCCAGGAAATTACTGATGAGTGCCTTACTCAGCCCTTATACCGGAGGA 3273  
QY 2671 actaaatgaagaatttggaccctcggggccatataaagacagaagttgaagataat 2730  
Db 3274 ACTAAATGAACATTTGGAGCTCCGGGCCCATATATTAAGACAGAGTTGAAGATAATAT 3333  
QY 2731 catgttaacttcaagaatcagagcctcctgcctattccttcatcttactatctatctt 2790  
Db 3334 CATGTAACTTTCAGAATATCAGGCTCTCGTCCCTATTCTTCTATTCTTACCTTATTTTC 3393  
QY 2791 ttaatgaagaatcagagagcagaagcagaacactagaaaaaactltgltcaagcctaatga 2850  
Db 3394 TTATAGAGAAATCAGAGGCAAGGAGCAAGACCTAGAAAAAATTGTGCAAGCCTAATGA 3453  
QY 2851 aaccaaacttacttcttggaaagtgcacaatcataatggcccaataaagattga 2910  
Db 3454 AACCAAAACTTACTTTTGGAAAGTGCAACATCATATGACCCCACTAAAGATGAGTTTGA 3513  
QY 2911 ctgcaaaagcctgggacttattctctgtatgttgaacttgaaagaagatgacactaagcct 2970  
Db 3514 CTGCAAAAGCCGGGCTTATTCTCTGATGTTGACCTGAAAAAATGTGCCTAGGCTT 3573  
QY 2971 gattggacccctctgtctgcccacactaacacactgaacccctgcatggyagaagaat 3030  
Db 3574 GATTGAGACCCCTTCTGCTGCCACACTAACACACTGMAACCTGCTCATGGAGACAAGT 3633  
QY 3031 gacagttacagaatattgtctcgttttccacaactcttgaagagccaaaagctggtactc 3090  
Db 3634 GACAGTACAGGAATTTCTCTGTTTTCACCATCTTTGATGAGACCAAAACCTGTGACTT 3693  
QY 3091 caatgaanaataatgaagaacactcagaggtccctgtcaatataatcagaatggaagatccac 3150  
Db 3694 CACTGAAAATTTGAAAAGAACTCAGGGCTCCCTGCAATATCCAGATGGAGATGCCAC 3733  
QY 3151 tttaaagaagaatatcgtctccatgcaatcaatggtctacaaatgaataacactg 3210  
|||||

Db 3754 TTTTAAAGCAATTAATCCCTTCCATGCATCAATGAGGCTACATAATGATACACTACCTGG 3813  
QY 3211 cttagtaatggctcagaatcgaatgaatcgaatggtatcgtcgaagatggccaagaatga 3270  
Db 3814 CTTAGTAATGGCTCAGGATCAAAAGGATTGATGATGTCTCAGCATGGGACCAATGA 3873  
QY 3271 aaacatccatctatcttcaatcagtgagacatglttcaactgtaaaaaaagaagaatga 3330  
Db 3874 AAACATCATCTATTTCACTTTTCACTGAGCATGTGTCTCTACGAAAAAAGAGAGATA 3393  
QY 3331 taaaatggactgtacaatcctatccagtglttltgaagacagtgaaatglttaccatc 3390  
Db 3934 TAAATGGCACTGTACAAATCTCTATCCAGTGTGTTTGTAGACAGTGAATGTTACATC 3993  
QY 3391 caaagcttgaaatttggcgggtgggaatgcttatttggagagacatctacatgcttggatag 3450  
Db 3994 CAAGCTGGAATTTGGCGGGGGAATGCTTATTTGGCAGAGATCTACATGCTGGGATAG 4053  
QY 3451 cacactttctgtgtgtaacaagaatgaatgltcagaactccctgggaatggtcttctgaga 3510  
Db 4054 CACACTTTTCTGCTGTACAGCAATAGTGTACAGACTCCCTGGGAATGGCTTCTGGACA 4113  
QY 3511 catlagagaatttcaagattacagcttcaagaacaatataatgaagltggcccaagctggc 3570  
Db 4114 CATTAGAGATTTTTCAGATTACAGCTTCAGACAATATGACAGTGGGCCCAAAAGCTGGC 4173  
QY 3571 cagacttcatattcccgatcaatcaatgactgtgagacccaagagcccttcttggat 3630  
Db 4174 CAGACTTCATTATTTCCGATTAATCAATGACCTGGAGCCAAAGAGCCCTTTCTTGAT 4233  
QY 3631 caagctgagatctgttgcacaaatgatatatacaaggcaatcaagaacccaggtgtccgcga 3690  
Db 4234 CAAGTGGATGTTGGACCAATGATATTATTCAGCGCATCAAGACCCAGGCTGCCGCTCA 4293  
QY 3691 gaagltccagcctcactacatcctcaglttatcaatcatgatatglttgaatggaaga 3750  
Db 4294 GAACTTCTCCAGCCTCTCATCTCTCAGTTATCATCATGATAGTGTGATGGAGAA 4353  
QY 3751 gtgcagagacttatcgaagaatctcacactggaaccttaatglttcttcttggcaatgaga 3810  
Db 4354 GTGGCAGACTTATGAGGAAATTCACCTGGAACCTTAATGATGCTTTTGGCAATGTGA 4413  
QY 3811 ttcacttgggaataaacaataatlttlaacccctcaatlttgcgtacatccogtt 3870  
Db 4414 TTCACTGGGATTAACCAATATTTTAAACCTCCCAATTTTGTCTGCATACATCCGTTT 4473  
QY 3871 gcaaccaactcatatagacttcgagcactcttcogaltgagatgagtggtctgattc 3930  
Db 4474 GCACCAACTCATTAATAGCATTTCCAGGACTCTTCGATGGAGTTGATGGGCTGTGATTT 4533  
QY 3931 aaatagttgcagcatgcatattgggaatggagatgaagaatataatcagaatgacattac 3990  
Db 4534 AAATAGTTGCAGCATGCGCATTTGGAAATGGAAGATAAACAATATCAGATGACAGATTAC 4593  
QY 3991 tgtctcatcacttacttacaatataatgttgcacactggtctccttcaaaagctcgaactca 4050  
Db 4594 TGCCTTACCTCACTTTACCAATATGTTTGGCACTGGTCTCTTCAAAAGCTCCGACTTCA 4653  
QY 4051 cctccaagggaggaatattgcttggagacactcagtggaataatccaaaagatggcttga 4110  
Db 4654 CCTCCAAGGGAGGAGTAATGCTGTGAGACCTCAGGTGAATATCCAAAAGAGTGGCTGCA 4713  
QY 4111 agtggacttccagaagaagaatgaagaatcagaaggttaactatcgaaggttaaaatctctc 4170  
Db 4714 AGTGACTTCCAGAAAGCAAAAGAAAGTACAGGAGTAACTACAGGGAGTAAATCTCT 4773  
QY 4171 gcttaccagcatgatatgtaagaggttcccatatccagcagatcaagaatggtccatcagtg 4230  
Db 4774 GCTTACCAAGCATGTATGTGAAGAGAGTTCTTATCTCCAGCATGAAGTGGCATCAAGT 4833  
QY 4231 gactccttcttcaagaatggcaagtaagaagtttccaaaggaatcaagctccttccac 4290  
Db 4834 GACTCTCTTTTTCAGAAATGCAAAAGTAAAGTTTTTCAGGGAATCAAGACTCCTTAC 4893  
|||||



QY	4291	acgtttgtggaactctctctgagaccacccggttactgacactcgcgttacttgaattcaacccca	4350
Db	4894	accttggtggaactctctgagaccacccggttactgacactcgcgttacttgaattcaacccca	4953
QY	4351	gagtttggtgacccaagattgcacctgagatlgagagttctctggtcctgcaagacacagagact	4410
Db	4954	gagtttggtgacccaagattgcacctgagagatlgagagttctctggtcctgcaagacacagagact	5013
QY	4411	ctacttga 4417	
Db	5014	ctacttga 5020	
RESULT	8		
AR071306			
LOCUS	AR071306	5035 bp	DNA
DEFINITION	Sequence 1	from patent US 5910481.	PAT
ACCESSION	AR071306		18-FEB-2000
VERSION	AR071306.1	GI:7222194	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	Unclassified.		
AUTHORS	1 (bases 1 to 5035)		
TITLE	Voorberg, J.J.		
JOURNAL	Hybrid proteins with modified activity		
FEATURES	Patent: US 5910481-A 1 08-JUN-1999;		
source	Location/Qualifiers		
	1..5035		

BASE COUNT	1484	a	1127	c	1110	g	1314	t
ORIGIN	/organism="unknown"							
Query Match	84.8%	Score	3743.8	DB	9	Length	5035	
Best Local Similarity	87.5%	Pred. No.	0					
Matches 4365: Conservative	0	Mismatches	52			Indels	570	Gaps
							2	

QY	1	catcgaataagagctctccacgtctctctctctgtgaccttggacttgcattagtc	60
QY	11		
Db	34	CATGGAAATAGAGCTCTCCACTGCTCTCTTCTGTGCTTTGGAGATTGCTTATGTC	93
QY	61	caaccagaatactactactctgggtgcagtggaactctcatggaactatctgcaagtgatct	120
Db	94	CACCAGAAGATACACCTCGGGGTGAGTGGAACCTGTCAATGGACATATATGCAAGTGATCT	153
QY	121	cgtggaagctgacctgtgagcgaagattctccctcagaagtgcgaatactttccattcaa	180
Db	154	CGGTGAGCTGCTGTGGACGCAAAATTTCCCTCCTAGATGTCGCAAAATCTTTTCATTCAA	213
QY	181	caacctagctgtgtacaanaaagactctgtgtgtagaattcaagtgltcaacttccaacat	240
Db	214	CACCTCAAGTCTGTACAAAAAGACTCTGTTGTGTAGATTACGGATCACCTTTTCAACAT	273
QY	241	cgcctaaagccaagsgccacctgtagtgggtctgtcagtgccctacatctcaagcttgagttta	300
Db	274	CGCTTAAGCCAAAGGCGACCCCTGAGTGGCTGTCTGATAGTCACCATCCAGGCTCAGGTTTA	333
QY	301	tgtataagctgtgtcaacttaacacttaagaacaatggcttccatctcgttcagttcatgtgt	360
Db	334	TGATATCAAGTGTCTATTACACTTAAGAACATAGGCTTCCCATCTGTCTCAGTCTTCAATGCTGT	393
QY	361	tgtgtatctactacttgaaagctctctgtaggaagctgaatgatgatcatcagaaccgaagaag	420
Db	394	TGGTGTATCTTACCTGGAAGCTTGTGAGGGAGCTGGAATATGATGATCAGACCGATCAAG	453
QY	421	ggagaaagaagaatgaataagctctccctctgtgtggaagcacaagatgctcggagaagcttc	480
Db	454	GGAGAAAGCAAAATATATAAGCTTCTCCCTGGTGGAGGCACTACATATATCTGGCAGGGTCT	513
QY	481	gaaagagaatgttccaatgtgctcttgaccactgtgcttactactcatatatttctca	540

Db	514	GAAGAGAAATGGTCCAAATGGGCTCTGACCCACATGTGGCTTACTACTCATATCTTTTCGA	573
OY	541	tgtggaaccttgtaaaagaacttgaatcctaaggctcctaattggaagccctactagtataga	600
Db	574	TGTGGACCTCGTAAAGACTTGAATTTCAGGCCCTCATTTGGAGCCCTCTAGTATGTAGAGA	633
OY	601	aggagatcttgcgaaggaaaagaacagagaccttgacaaatctatactactcttttgcgt	660
Db	634	AGGGAGTGTGGCCAAAGGAAAAAGACACAGACCTTGCAAAATTATCTACTTTTGGCTGT	693
OY	661	atttgaatgaagggaagaatttggcactcagaacaagaagactcccttgaatgcagataagga	720
Db	694	ATTTCATGAAGGGAAGAAGTTGGCACTCAGAAACAAAGAACTCCTTGATTCAGAGTAAGGA	753
OY	721	tgctgcacctctctggggccctggccctaaatagcaacaaagtcgaatggtgatgaagaagtc	780
Db	754	TGCTGCATCTCTCTGGGCTGGGCTTAAATGACACAGTCATAGGTATGTAAACAGGTC	813
OY	781	ctctgcgaagctctgaattgagtgcccaagaaatcagtcatacttggcatgtatgtgaatbvg	840
Db	814	TCTGCGAGGCTTGATTTGGATGGATGCCACAGGAATCATGTCATTTGGCAGTGTATTGGAA	873
OY	841	caccaactctgaagttgcactcoataatctctcgaaagttcaacaattcttgytgaagacca	900
Db	874	CACCACCTCTGAAGTGCACATCAATATTCCTCGAAGGTCAACACTTCTGTGAGGACCA	933
OY	901	tcgcgaagccctcccttggaaatctcgccaataactctctactgctgcctcaaaaccttgat	960
Db	934	TGCGCAGGCTCTTTGGAAATCTGCCAATATCTTTCTTACTGCTCAACACCTTTGAT	993
OY	961	ggaccttggaaagtttctactggttttgcaatactctcccaaccaacatgatatgcaatga	1020
Db	994	GGACCTTGGAAGTTTCTACTGTTTGTCTATCTCTCCACCAACATGATGATGATGGA	1053
OY	1021	agcttatgtcaaaatgaacagcgtctcgcaagggaaccccaactcagaatgaataaataaga	1080
Db	1054	AGCTTATGTCAAATGACAGCTGTCCAGAGGAACCCCACTACGATGAATAAAATATAGA	1113
OY	1081	agaagcggaaagacatgatgatctacttgaattcttgaatgtgaatgtgtgtcaagtttga	1140
Db	1114	AGAAGCGGAAGACTATGATGATGATCTTACTGATTCGAAATGGATGTGTGAGTTTGA	1173
OY	1141	tgaatgaacaactctccctctcttataccaatcttgcctaagttgcgaagaagcatccctaaac	1200
Db	1174	TGATGACAACTCTCTCTCTTATACCAAAATTTGCTCAGTTGCCAACAACATCTCPAAAC	1233
OY	1201	cttggatcaatcacaatctgcctcgtgaagaggagacttggactatgctcccttgcctctgc	1260
Db	1234	TTGGGTACATTACATTTGCTGCTGTAAGAGAGACTGGACTATGCTCCTTATGCTCTGC	1299
OY	1261	ccccgatgaacgaagtataaagtcacaattttgaacaatbgyccctcagcggaattgttag	1320
Db	1294	CCCCGATGACGAAGTTATATAAAGTCAATATTTTGAACAATGGCCCTCACCGGATTTGAG	1353
OY	1321	gaagtacaaaaaagtcagattttagcatatacaagaatgaagactttaaagctcgtgaagc	1380
Db	1354	GAAGTACAAAATAACTCCGATTATATGGATACACAGATGAACCTTTAAAGCTGTGAAGC	1413
OY	1381	tattcagcatgaatcaggaactcttgggaaccttacttctatgtgggaagtlttggacaacact	1440
Db	1414	TATTACAGATGAATCAGAGATCTTTGGGACCTTACTTTATGGGAAAGTTTGGACACACT	1473
OY	1441	gttgaatatataatlaagaatcaagcaagcagacataataactctaacctccacggaataac	1500
Db	1474	GTTGATTATATTTTAAAGATATCAGACAGACACCATTTAACTCTACCTCCACGAGATATAC	1533
OY	1501	tgaatgctgcctcttgaattcacaaggagatctacaaaagglttaaaacatttgaagaattc	1560
Db	1534	TGATGTCCGCTCTTTGTATTTCAAGAGCAATTTACCAAAAGGTGTAACATTTTGAAGATTT	1593
OY	1561	tccaactctgcagaggagaaatlctcaataataatggaacagtgcagtgtatagaatltggcc	1620
Db	1594	TCCAAATTTCTGCAGAGAAATATTTCAAAATTTAATGTGAACAGTACGTGTGAAGATGGGCC	1653



QY 1621 aactaaatcaagatcctcgctgacctgacccgcctattacttaagtttgcgttaatatgagag 1680  
|||||  
Db 1654 AACTAAATCAGATCTCTGGTGCCCTGACCCGCTATTACTTAGTTTCGTTAATATGGAGAG 1713  
QY 1681 agatctagctcaagactcatttgccctctctcatcctatcgtctacaaagaatcgtatga 1740  
|||||  
Db 1714 AGATCTAGCTTCAAGACTCANTGGCCCTCTCTCATCTGCTGACAAAGATCTGTAGATCA 1773  
QY 1741 aagaggaaccagataatgtcagacaaaggaaatgcatccctgtcttcgtlatattga 1800  
|||||  
Db 1774 AAGGAGAAACCGATATATGTCAGACAAAGAGAAATGATCATCTGTTTCTGTGATTTGATGA 1833  
QY 1801 gaaccgaagctggttacctcacaagaatatatacaacgctttctcccaatccagctgag 1860  
|||||  
Db 1834 GAACCGAAGCTGGTACTCTACAGAGAAATATACAGCGTTTCTCCCAATCCAGCTGGAGT 1893  
QY 1861 gcaagcttgagatccagaagttccaaagcctccaacatcatgacagacatcaatggtat 1920  
|||||  
Db 1894 GCAGCTTGAAGATCCAGATGTCGAAGCCTCCAAACATCATGACAGCATATGGCTATGT 1953  
QY 1921 ttctgataagtttgcaagttgtcagttgtttgtcatgaggtgagcatcctgtaactc 1980  
|||||  
Db 1954 TTTTGAATAGTTTGCAAGTTGTCAGTTGTTGCATGAGGTGGCATATGTAACATTTCTAG 2013  
QY 1981 catgtgagcacaagactgacttcccttctgtcttctctctcgtgatatcccttcaaa 2040  
|||||  
Db 2014 CATTTGGACACAGACTACTCTCTTCTGTCTTCTTCTCTGAGATATACCTTCAACACAA 2073  
QY 2041 aatgctatagaagacacactcaacctatcccatctcagagaaactgtctcatagtc 2100  
|||||  
Db 2074 AATGCTTATGAAGACACACTCACCTTATTCCTCATCTTCAGAGAAACTGTCTTCAAGTC 2133  
QY 2101 gatygaaaccaggtctatagatctcggggtgcacaactcagaacttcggaaacagg 2160  
|||||  
Db 2134 GATGGAACCCAGGCTATGAGATTCGGGGTGCACACACTGCAACTTTCGGAACAGAG 2193  
QY 2161 catgacgccttactgaagttctctagttgtgacaagaacagtgatattatcaggga 2220  
|||||  
Db 2194 CATACCGCCTTACTGTAGAGGTTTCTAGTGTATTCAGAGGGGAGAGGCGCGACTA 2253  
QY 2221 caattatgaagatatctca-----gcatactgtcgtagtaaaacaatgc 2265  
| | | | |  
Db 2254 TCTGCACTGAGAGATATTCAGTGAAGAGACGACTACATGACATGCGCACTGCT 2313  
QY 2266 catggaaccagaagctctcccaagaattcaagaacccctagaactcggcaaaagaatt 2325  
|||||  
Db 2314 GATTGAACCAAGAGCTTCTCCAGAAATCAAGACACCCCTAGCACTAGGCAAAAAGCAATT 2373  
QY 2326 taatgacac----- 2334  
|||||  
Db 2374 TAAATGCCACCAATTCACAGAAATGACATAGAGAACGACCTTGTTGACACAG 2433  
QY 2335 ----- 2334  
Db 2434 AACACATATGCTTAATATACAAATGTCTCTAGTATTTGTTGATGCTTGGCACA 2493  
QY 2335 ----- 2334  
Db 2494 GAGTCTACTCACANGGGCTATCTTATCTGATCTCCAGAAAGCCAAATATGAGACTTT 2553  
QY 2335 ----- 2334  
Db 2554 TTCTGATGATCCATCACTTGAGACAAATAGACAGTAATAACAGCTGTCTGAATGACACA 2613  
QY 2335 ----- 2334  
Db 2614 CTTCAGGCAACAGCTCCATCAGTGGGACATGTATTTACCCCTGATGAGGCTTCCA 2673  
QY 2335 ----- 2334  
Db 2674 ATTAAGATTAATGAGAAACTGGGACACTGCAGATCTCTTGTGATTAACCACTA 2733

QY 2335 ----- 2334  
Db 2734 TGTACTCAGATACCAAAAGAGAGTGGAATCCCAAGAGAAAGTACACAGAAAAACAGC 2793  
QY 2335 ----- 2334  
Db 2794 TTTTAAGAAAAAGATACCAATTTTGTCCCTGAACGCTTGTGAAGCAATCATGCAATAGC 2853  
QY 2335 ----- 2334  
Db 2854 AGCAATTAATGAGGACAAAATAAGCCGGAATATGAAATGACCTGGCCAAAGAGGTAG 2913  
QY 2335 -----cccaacagctcttgaacgcacatacaagggaaataac 2370  
|||||  
Db 2914 GACTGAAGGCTGTGCTCTCAAAACCACAGTCTTAAACGCCATACAGGGAAATATAC 2973  
QY 2371 tctgtactctcttcagtcagatcaagaagaattgactatgatatgatalccalatcagttga 2430  
|||||  
Db 2974 TCGTACTACTCTTCAGTCAGATCAAGAGAAATGACTATGATGATACCATATCAGTTGA 3033  
QY 2431 aatgaaagaagaagatttgacattatgatatgagatgaaatcagaagccccgcagctc 2490  
|||||  
Db 3034 AATGAAGAGGAAGATTTTACATTTATGATGAGGATGAAATACAGCCCCGACACTT 3093  
QY 2491 tcaaaagaaaacacgacactatttctatgtcagtgagagagctctgagattagagat 2550  
|||||  
Db 3094 TCMAAAGAAACACGACACTATTTATTTGCTGCAGTGGAGAGGCTCTGGGATTTTGGGAT 3153  
QY 2551 gagtaagctcccaacatgtctcctaagaaacaggtcacaagtgtgcagtgctccctcagttcaa 2610  
|||||  
Db 3154 GAGTAGCTCCCAACATGTCTTAAAGAAACAGGGCTCAGAGTGGGAGTGCTCCCTCAGTTCAA 3213  
QY 2611 gaaagttgcttcacgaagaatttactgattgctccttactcagcccttatacgttgaga 2670  
|||||  
Db 3214 GAAAGTGTTCGAGGAATTTACTGATGTGCTCCTTTACTCAGGCCCTTATACCTGTGAGA 3273  
QY 2671 actaaatgaacatttggagctccctcggggccatatataagaagcagaagtltgaagaataat 2730  
|||||  
Db 3274 ACTTAATGAACATTTTGGGACTCCTGGGGCCATATATAGACGACAGCTTGAAGATTAATAT 3333  
QY 2731 catgttactctcagaatacaggcctctgtcccatctctctctatctatctagccttattc 2790  
|||||  
Db 3334 CATGTAACTTTTCAGAAATCAGGCGCTCGTCCCTATTCTTCTATCTAGCTTATTTTC 3393  
QY 2791 ttatgaggaagatcagagggcagaagagagaacttagaanaaaacttgtcaagcctaata 2850  
|||||  
Db 3394 TTATGAGGAAGATCAGAGGAGAGGAGGAGAACCTTAGAAAAAATCTTGTCAAGCCTAATGA 3453  
QY 2851 aaccaaaacttacttttggaaagtgcacatacatatgaccaccaactaaagatggttga 2910  
|||||  
Db 3454 AACCAAAACTTACTTTTGGAAAGTGCAAACATCATATGCGACCCCTAAGATGAGTTGA 3513  
QY 2911 ctgcnaagccttggcttattctctctgtatgtgaccttgcgnaaaagatgtgactcaggct 2970  
|||||  
Db 3514 CTGCAAAAGCCTGGGCTTATTTCTCTGATGTGACCTGGAAGAAAGATGTGACACTAGGCT 3573  
QY 2971 gatttgacccctctgtctgtcgcacactaaacaactgaacccgtctatgtggagacaagt 3030  
|||||  
Db 3574 GATTGACCCCTTCTGTGTCGCCACACTTAACACTGAACCTGTGCTATGTGGAGACAAAGT 3633  
QY 3031 gacagttacaggaatttgctgttcttccacatcttggatggagaccnaagctgtgtactc 3090  
|||||  
Db 3634 GACAGTACAGGAATTTGCTCTGTTTTACACATCTTTGATGAGACCAAAAGCTGTGACTT 3693  
QY 3091 cactgaaatatlgnaaagaacatgcagggctccctgtcaatalccogaatggaatccac 3150  
|||||  
Db 3694 CACTGAATAATATGGAAGAAATGCAAGGCTCCCTGCAATATCAGATGGAAGATCCAC 3753  
QY 3151 tttaaagaagaatatatgccttcatatgcaatcaatggttaatatatgatactactctg 3210  
|||||  
Db 3754 TTTTAAAGAAATTAATTCCTTCATGCAATATGCGTATCATATGATGATACACTACTGG 3813  
QY 3211 cttagtaatgtcgaagatcaaaagatctgatgtatctcctgaacatggtgcagcaatga 3270







Db 543 CACGTGACCGCTGTGAAGACCTGAATTCGGGCTCATTTGAGGCCCTGTGGTTGTAGA 602  
Qy 599 gaagaggtctggccaagaaagacacagacctgtgcaaatltataactctttgtc 658  
Db 603 GAAAGGAGTCTGACCAAGAAAGGAGCCAGAACCTGCACGAATTTGACTACTTTTGGCT 662  
Qy 659 gtattcatgaaggaaagtctggcactcagaacaagaacctcctgtatgcagatag 718  
Db 663 GCTCTTATGAAGGAAAAAGTGGCACTCAGCAAGAAATGATCTCTTGACACGGGCCATG 722  
Qy 719 gatgtcatctgtctcggtcctggcctaaatgtcacacagltcaatggtltaatgaag 778  
Db 723 GATCCCGACCTGCGCAGGCGCCAGCCTGCATGACACAGATGATGCTATGTCACAGG 782  
Qy 779 tctctgcaggtctgtatgtatgtccacagaaatcagtcattgtgcatgtgatgtg 838  
Db 783 TCTCTGCGAGGTGTGATCGGATGTCTATGAAGAAATCAGTCTACTGGCAGTATTGGAAAG 842  
Qy 839 ggcacacactcctgaagtgaactcaatatctcctgaaggtccacacattctctgtgaggaac 898  
Db 843 GGCACACCGCCGGAAAGTGCATCTATTTTCTGAAGGCCACACGCTTCTGCTGAGGCAC 902  
Qy 899 catcgcaagcgctcctgtggaatctcgccaataacttccctactgtctcaaacctctg 958  
Db 903 CATGCCAGGCTCTTGAGATCTCGCACTACTACTTCCCTCACTGCTCAGACATTCCTCG 962  
Qy 959 atggaacttggacagttctactgttttgttcatatcctctcccaacaatgatgtgcatg 1018  
Db 963 ATGACCTTGGCCGCTTCACTGTTTGTGTCTATCTCTCCGCCACACATGGTGGCATG 1022  
Qy 1019 gaactcatltcaaaagtacagctgtccagaggaaccccaactagaatgaaaaaat 1078  
Db 1023 GAGGCTACGTCAGATGAAAGCTGCGCCGAGAGGCCCAAGCTGCGGAGAAAGCTGAT 1082  
Qy 1079 gaagaaacggaagactatgtatgtacttacttacttctbtaagatgtgtgtcaagttt 1138  
Db 1083 GAAGA---GGAAAGTATGATGACAAATTTGACAGACTCGGACATGAGAGCATGCCAAA 1139  
Qy 1139 gatgatgaacaactcctccttccattatccaaatcogtcaagttgcagaagacatcctaaa 1198  
Db 1140 GATGGTACGACGCTGTCTCCCTTATTCAAATCGCGTGGTCAAGAAAGCATGCCAAA 1199  
Qy 1199 acttggatcatlatactgtctgtcgaagagagactgggactatgtctcccttactctc 1258  
Db 1200 ACCCGGTGACACTACATCTCTGCAGAGAGAGAGACTGGGACTACGCCGCCGCTCCC 1259  
Qy 1259 gcccccgatgacagaagttaataaagltcaatlatttgaacaatggtccctcagcgatgtgt 1318  
Db 1260 AGCCCGAGTGACAGAACTTATTAAGTCTCTACTTGAAACAGTGTCTCTCAGCGAATTGGT 1319  
Qy 1319 aggaagttaacaaaaaagtcgaattatgtgcatacacagatgaacaacttgaagctgtgaa 1378  
Db 1320 AGGAAATTAACAAAAAGCTCATTCGTCTGCTACCGATGTATACATTATGAGACTCGTAAA 1379  
Qy 1379 gtattcagcatgaatcaggaatctgtggaaccttacttattatgsgaagtgtgagaaca 1438  
Db 1380 GCTATTCCGATGATGAGAAATCTGTGGACCTTTACTTTATGAGAGAAAGTTGGAGACACA 1439  
Qy 1439 cgttgtatataattaaagatcaagcaagacacataataactltaacctcagcgatc 1498  
Db 1440 CTTTGTATATATTAAAGATTAAGAGCGAGCCGACATATACATCTACCTCATGGAATC 1499  
Qy 1499 actgatgtccgtctctgttctcaagagatltacaaaagtgttaaacacttgaagat 1558  
Db 1500 ACTGATATCAGCGCTTGGACCCAGGAGACTTCTAAAAGGTTGGAAACTTTTGAAGAC 1559  
Qy 1559 ttccaattctgcagaagaaatatacaaatataaagtgcagtgactgtgagaagatgag 1618  
Db 1560 ATGCAATTTCTGCCAGAGAGACTTTCAAGTATTAATGAGACAGTGTGGAAGATGGG 1619  
Qy 1619 ccaactaatcagatcctcgtgtccgtgaccgctatctactagtttcgttaatatgag 1678

Db 1620 CCAACCAAGTCCGATCCCTGGTGGCTGACCCGCTACTACTGAGCTTCATTAATCTAGAG 1679  
Qy 1679 agaatctagcttcaaggaactatgtggccctctcctactctgtcacaagaatcgtatg 1738  
Db 1680 AAGATCTGGCTTGGAGCTCATTTGGCCCTCTCTCTCATCTGCTACAAAGATCTGTAGAC 1739  
Qy 1739 caaaggaagaacacagaataatgtcagacaagaagatbtcaatccgtlttctgtatgtat 1798  
Db 1740 CAAGAGAGAAACCGATGATGTGACACAGAGAAAGCTCATCTCGTTTCTGTATTCGAT 1799  
Qy 1799 gagaacccaagctgtgtactcctcaagaaatatacaacgcttctcccaatccaagctgga 1858  
Db 1800 GAGATCAAGCTGTGTACCTCGCAGAGAAATTCAGCGCTCTCCCTCCCAATCCGATGGA 1859  
Qy 1859 gtgcagcttgaagatccaaggttccaagctccaactcactgaacagatcaatagctat 1918  
Db 1860 TTACAGCCCCAGATCCAGAGTTCCAAGCTTCTAACATCATGACACACATCAAGGCTAT 1919  
Qy 1919 gttttgatagttgcaggtgtcagttgttgcataaggtgcatactgtatattccta 1978  
Db 1920 GTTTTATATAGCTTGCAGCTGTGCGTTTGTGACAGAGGTGGCATCTGTATATTCTA 1979  
Qy 1979 agaatggaacaaagactgtacttcttctgtcttcttctctgtatataacttcaaacac 2038  
Db 1980 AGTGTGAGCACAGACGAGCTTCTCTCCGCTCTTCTGTGCTACACTTCAACAC 2039  
Qy 2039 aaaaatgtcatgaagacaacttaccctattcccatlctcagaagaactgttcttatg 2098  
Db 2040 AAAATGCTATGAAAGACACACTCACCTGTCTCCCTTCAGAGAGAAAGGCTTTATG 2099  
Qy 2099 tcatgaagaaccccaagttcatatgatcttgggtgtgcacaactagaacttgcagaaga 2158  
Db 2100 TCAATGAAAAACCGAGTCTCTGAGTCTTAAGGTGTCACCAACTCAGACTTGCAGAACAGA 2159  
Qy 2159 gcaatgcgccttactgaaggttctcagttgtgtgacaagaacactgtgtatattacagag 2218  
Db 2160 GGGATGACACCTTACTGAAGGTGTATGTTGTACAGGAGCATGTGATATTATATGAC 2219  
Qy 2219 gacagttatgaagatatttcaagcacttctgtcgtatgaataaacaatgtccattgaaccaaga 2278  
Db 2220 AACCTTATGAAGATATTCAGAGCTTCTGTGATGGAAGAAATGTCATTTGAACCCAGA 2279  
Qy 2279 agcttcccagaatltcaagaacccctagcactagsgaaaaagcaatttaatgccaacca 2338  
Db 2280 ----- 2279  
Qy 2339 ccagttctgaacgcacatcaacgynaataactcgtactactccttcagtcagatcaagag 2398  
Db 2280 -----GACATTAAGCCTTCTACTTTTCAGCCGAGAGAGAC 2315  
Qy 2399 gaaattgactatgatgtatccatatcagttgtgaatgtgaagaagaagatttgcattat 2458  
Db 2316 AAAATGACATATGATGATATCTTCTCAACGTGAACGAGGAGAGAAATTTGACATTTAC 2375  
Qy 2459 gatgaagatgaaaatcagagccccgcagcttccaagaagaacaacacacacttattat 2518  
Db 2376 GGTGAGATGAATAATCGACACCTCGCAGCTTTCAGAAAGAAACCCGACACTATTTTATT 2435  
Qy 2519 gctgcagttgagaagctcctggtattatgagatgagtagtctcccaactgttctaagaac 2578  
Db 2436 GCTGCGGTGAGCAGCTCTGGGATTTACGGGATGAGCAATTCCTCCCGCGGCTTAAGAAAC 2495  
Qy 2579 aggtgtcaaggtgtgcagttgtccctcagttcaagaagaagtgttcttccaagaaattcagat 2638  
Db 2496 AGGGCTCAGAACGAGAGGTGCTCGGTTTCAAAAGGTGTCTTCGGGAATTTGCTGAC 2555  
Qy 2639 ggtccttactacagccctataccgtgtgagaactaaataaactgtgagccctgtgag 2698  
Db 2556 GGCTCTTACGACGACCGGTGTACCGGGGGAACTCACAAACACTTTGGGCTCTTGGGA 2615  
Qy 2699 ccalataaagacagaaglttgaagataatacatgtgtaacttccagaaatcagccct 2758  
Db 2616 CCTACATCAGACGGAAATTTGAAGACATCATGTGTACTTTCAAAAACCGGCGCT 2675



QY 2759 cgtccctattctcttacttactagccttattcttaaggaagatcagaagcagagca 2818  
|||||  
Db 2676 CGTCCCTATTCTCTTACTACTGAGCCTTATTCTTATCCGATGATCAGAGCAAGGGGA 2735  
QY 2819 gaacctgaaaaaaccttctgaagcctaaatgaacaaactactcttttggaaatgcaa 2878  
|||||  
Db 2736 GAACCTGACCAACTTCGTCCAGCAATGAAACCAAGAACTTCTTGGAAATGCGAG 2795  
QY 2879 catcatatggaccocactaaagatbaggtttgaatgcaagcctgggcttatctctgat 2938  
|||||  
Db 2796 CATCACATGGACCCACAGAAAGAGAGATTGACTGCAAAAGCCTGGGCTACTTTTGTGAT 2855  
QY 2939 gttagacctgaaaaaagatgtgcaactcaggccttgatctggaccctctgtctgcacact 2998  
|||||  
Db 2856 GTTAGACCTGGAAAAAGATGTGCACACTCAGGCTTGATGGGCCCTCTGTGCTGCCGCCG 2915  
QY 2999 aacacactgaaacctgtctcatatggagacagatgacagatcaggaatttgctctgttttc 3058  
|||||  
Db 2916 AACACCTGAACGCTGCTCACGGTAGACAACTGACCGTGAAGATTTGCTGTGTTTTTC 2975  
QY 3059 accactcttgatgagaccacaaagctggtacttcaacttgaataatgaaagaaactgcag 3118  
|||||  
Db 2976 ACTATTTTGTATGAGACAAAGAGCTGTGACTTCACTGAATGTGGAAAGAACTGCGG 3035  
QY 3119 gtccctgcaatataccagatggaagatcccactttaagaagaattatcgcttccatgca 3178  
|||||  
Db 3036 GCCCCTGCGACCTGAGATGGAGAGCCCACTGTGAAGAAACTATTCGTTCCATGCA 3095  
QY 3179 atcaatgtctacataatgagatacactaactgtgcttaagttaatgctcagaatcagaat 3238  
|||||  
Db 3096 ATCAATGGCTATGTATGATGATGATACCTCCCTGGCTTATGATGCTCAATCAAAAGGATC 3155  
QY 3239 cgaatgcatctgctcagcatgagcaatgaaacatccatcttcatctcatcttgatgca 3298  
|||||  
Db 3156 CGATGGATCTGCTCAGCATGGGAGCAATGAATATCCATTCATTTTATTTAGCGGA 3215  
QY 3299 catgtgttcaactgtacgaaaaaagaagagatataaatgtgcaactgttacaatctatcca 3358  
|||||  
Db 3216 CACGTGTTCAGTGTACGGAAGAAAGAGAGATTAATAATGCGCTGTACAACTCTATCCG 3275  
QY 3359 ggtgtttttgagaaagatggaatgtttacatccaaagcttgaaatttggcgggtgaatgc 3418  
|||||  
Db 3276 GGTGTCTTTGAGACAGTGGAAATGCTTACCGTCAAGTTGGAATTTGGCGAATAAGATGC 3335  
QY 3419 cttaattggcagcatctacatgctcggagatgagcacaacttttctgtgttacagcaatag 3478  
|||||  
Db 3336 CTGATTGGCGAGCACTGTCAGAGCTGGGATGAGCACGACTTTCTTGCTGTACAGCAAGGAG 3395  
QY 3479 tgtcagaactcccttgggaatgtctcttgagacatltiagagatlttcagattacaagcttca 3538  
|||||  
Db 3396 TGTACAGCTCCACACTGGGAATGGCTTGTGAGCATTTAGATTTCACATACACACTTCA 3455  
QY 3539 ggaacataatgagcagtggtggcccaaaagctggcagacttcatctccggtatcaatcaat 3598  
|||||  
Db 3456 GGAAAGATATGAGACGTGGGCCCAAAAGCTGGCCAGACTTATTTATTTCCGGATCAATCAT 3515  
QY 3599 gctctggagacccaagagagcccttctcttggatcgaagtgtgctgtctgtggcccaatgatt 3658  
|||||  
Db 3516 GCTTGAGACCAAGAGATCCCACTCTCTGATCAAGGTGGATCTGTGTGGACCAATGATC 3575  
QY 3659 attcaaggcatcaagagccagaggtgtcccgltcagaagltctccagccttcaatctctcag 3718  
|||||  
Db 3576 ATTCACAGCATCATGATGAGCCAGGTGCCGTGCAGAAAGTTTCCAGGCTCTACATCTCCAG 3635  
QY 3719 ttatcatcatgtatagttctgtatggagaagaatgtgagcagacttcatcgaggaattccact 3778  
|||||  
Db 3636 TTTATCTATCATGTACATCTTTGACGGGAGAACTGGCAGATTAACCGAGGAGATTTCCAG 3695  
QY 3779 ggaacactaatgtgtctcttggcaatgtgataatcaltcgatgataaacaacataatttt 3838  
|||||  
Db 3696 GGCACCTTAATGCTTCTTTGGCAATGTGACGCAATCTGGGATTAACACATATTTTTT 3755

QY 3839 aacctccaattatgtctgatbaatccglttgacccaactcatatagcatctgcagc 3898  
|||||  
Db 3756 AACCTCCGATGTGTCGCTGGATCCGTTTGGACCCAAACACATTCAGATCCGAGC 3815  
QY 3899 actcttgcagtgagttatgttgctgtatgattaaatgtttgcaatgcatgttgagatg 3958  
|||||  
Db 3816 ACTCTTGCATGAGATTTATGAGTGTGATTTAAACATTTGACACATGCCCCGGGATG 3875  
QY 3959 gagaataagcaatatacagatgcaagatcttacttcatcttacttacaatatttt 4018  
|||||  
Db 3876 CAGATTAAGCATATCAGACTCAAGATTCAGGCGCTCTCTCCACATCAGCAATATTTT 3935  
QY 4019 gccacctgtctcttccaagaagctgcagcttcaactccaagggaggaatgctctgaga 4078  
|||||  
Db 3936 GCCACCTGCTCTCTTCAACAAGCCGACTTCACTCCAGGAGGCGAGAAATGCTTGCGA 3995  
QY 4079 cctcagtggaataatccaaaagatgtgtgtgaatgtgacttccagaagacaatgaaatgc 4138  
|||||  
Db 3996 CCCCAGGAGACAGCGAGAGAGTGGCTGCAAGGTGACCTGCAAGAGCGTGAAGTTC 4055  
QY 4139 acaagatgataactcaagagagtaaatctctgtcttaccagacatgatagtgaaagatgc 4198  
|||||  
Db 4056 ACAGGCATCACACCCAGGCGGTGAAGTCCCTGCTCACAGACATGTATGAAAGACTTC 4115  
QY 4199 ctcatctcagaagatcaagatgtgcaatcagatggaactcttcttccagaatgcaagta 4258  
|||||  
Db 4116 CTCGTGTCAAGTACAGACAGCGCCGCTGAGCCCTGTGTTCTTCAAGAGCGCACACG 4175  
QY 4259 aaggttttccagggaatacgaagactcttcaactctgtgtggaactctctagaccaccg 4318  
|||||  
Db 4176 AAGGTTTTTCAGGGCAATCAGAGACTCTTCACCCCGGGTGAAGCTGTGGACCCCG 4235  
QY 4319 ttactgctcgtactcacttgcagatcgaatccccaagatgttggatgacacagatgtgcctgag 4378  
|||||  
Db 4236 CTGTTACAGCGCTACTAGATGATCCACCCAGAGCTGGGGGAGCAGCATTCGCCCTGAG 4295  
QY 4379 atgagatgtctgtgctgcagagcagaagcctctactgca 4417  
|||||  
Db 4296 CTCGAGTTCATGAGATGTGAGGACAGAGATCTCTACTGA 4334

RESULT 10  
ARI10041 4451 bp DNA PAT 14-FEB-2001  
LOCUS ARI10041  
DEFINITION Sequence 42 from patent US 6114148.  
ACCESSION ARI10041  
VERSION ARI10041.1 GI:12826317  
KEYWORDS  
SOURCE  
ORGANISM  
Unknown.  
REFERENCE  
1 (bases 1 to 4451)  
AUTHORS Seed,B. and Haas,J.  
TITLE High level expression of proteins  
JOURNAL Patent: US 6114148-A 42 05-SEP-2000.  
FEATURES  
source 1..4451  
BASE COUNT 972 a 1568 c 1234 g 677 t  
ORIGIN

Query Match 58.9%; Score 2603.8; DB 9; Length 4451;  
Best Local Similarity 74.8%; Pred. No. 0;  
Matches 3262; Conservative 0; Mismatches 1097; Indels 0; Gaps 0;

QY 59 gccaccagaagatactacttgggtgagtggaactgtcatatggaatgataatgaaatgcat 118  
|||||  
Db 84 GCCACCCCGCTACTACTGAGCGCGGTGAGAGCTGTCTGGGACTCATGACAGACGAG 143  
QY 119 ctggtgtgagcgtcgtgtggaagcaagaatttctctctagatgtgcaaaatctttctcatic 178  
|||||  
Db 144 CTGGGCGAGTCCCGGTGAGACCGCGCTTCCCGCGGTGCCAGAGAGACTTCCCTTC 203



OY	179	aacccctcaagtcgtgtacaaaaagactcgtttgtgaaattcaagtglttaaccttttcaac	238
Db	204	AACCAACGACGGGTGGTATCAAGAAAACCGTTTCGTGTGATACCGACACCACTGTTTAAAC	263
OY	229	atgcctaaagcaagaagccaccttgatggtgtcctcgtagctgcctacatccaggtcgaagtt	298
Db	264	ATTGCCAAGACCGCGGCCCTCTTGATGTGGCTTCTTGCCGCCACCATTCAGAGCCGAGGTG	323
OY	299	taigatacagtygtlcaattacttaacttaagaacatggtctccactcgttcagttcattgct	358
Db	324	TACGACACCGTGGTATCACCTCGAAGAACATGCGACGCCACCCGTGAGCTCGACGCC	383
OY	359	gttggtgtatccctactcggaaagcttctcggagctcgaatagatagatlcagaccagttaa	418
Db	384	GTGGGGCTGAGCTACTGTGAAGGCGACGAGGGGGCCGATGACGACGCCAGCTGCCAG	443
OY	419	agggagaagaagaatgataaagctctccctcgttggtaagcatalatgctcgaagtc	478
Db	444	CGCAGAAGGAGGAGCGACAGAGTGTTCGGGGGGGAGCCACACCTAGGTGGAGGTG	503
OY	479	ctgaagaagaatggtlccaatggtccctcgaaccactgtgccttactactataltcttct	538
Db	504	CTTAGAGAGAACCGCCCTATGGCCACGACCCCTGTGCTCACTCAACAGTACTCTGAGC	563
OY	539	catbgaccctcgttaaaagactgaaattcaagccctattggagccctcctagtatgata	598
Db	564	CACGTGGACCTGGTGGTAGGATCTGTAACACGCGGCTATGTGGCGCTGTGTGTGTGC	623
OY	599	gaaggagatcttggccaaggaagaagacagaccttgcacaaattatactaattttgct	658
Db	624	GAGGGCAGCTGGCCAAAGGAGAAACCCAGACCCTGCACAAATTCATCTCGTGTGCC	683
OY	659	gtatttgatgaagggaaaagtgtgcactcaagaacaagaactccttctgatgcagatag	718
Db	684	GTGTTCCACAGGGGGAAGACTGTGCACAGGAGACTTAAACAAAGCCTGATGAGGACGCG	743
OY	719	gatctcatctctcctcggcctcgtgccttaaatgacacagtcataatggtatgtlaaacag	778
Db	744	GAGCGCGCCAGCGCCCGCGCTGGCCCAAGATGTCAACCGTTAAGGGGTAGGTGAACGCC	803
OY	779	tctctgcgaagtcctgattgtgatagcacaggaataactagttatltgcatgttgatggaat	838
Db	804	AGCCTTGCCCGGCTGATCGGGCTCCACCGCAAGCGTGTACTGGACGTACTATCGGCATG	863
OY	839	ggacacactcctgaagtgcactcaatcattccctcogaagtcacacattccttcttggagac	898
Db	864	GGCAGCACCCCTAGGTGGCACACATCTTCTGTGAGGGGCCAACCTTCTGTGTGGCAGAC	923
OY	899	catgcgcagcgctccttggaaatctcgcacaataactttccttactgtctcaaacacttgc	958
Db	924	CACGCGCACGAGCCAGCCTGGGATCACCCCCACATCACCTTCTCACTAGCCAGACCTTGCTG	983
OY	959	atggaccttggagaagtttctactcgttttgcabartctcccaacaacatgatyagatg	1018
Db	984	ATTGACCTAGGCGCATTTCTCGTGTTCGCGACATTCAGCAGCCACCAAGACAGCGCATG	1043
OY	1019	gaagcttatgtlcaaaatlagacagctgtlccaagaggaaccccaactcagaatgaaaatat	1078
Db	1044	GAGGCTTAAGTGAAGGTGGAGACAGCTGCCCGGAGGAGACCCAGCTGCGATGAAGAACAC	1103
OY	1079	gaagaagcggaaagactatgatgtgattcttactgtattcgtaaatgtagtgtgtaagttt	1138
Db	1104	GAGGAGGCGCAGGACTACGACGACGACCTGACCCGAGAGGAGATGTGTTGTTGAGCTTC	1163
OY	1139	gatgatcaaacctctcttccttcaatccaattcgtctagtttggccaagaagatccaa	1198
Db	1164	GAGAGGACAAACAGCCCACTTCAATCCAGATTCGAGGCTGGCCAAAGACACCTTAAG	1223
OY	1199	acttggttacattacattgctcgtctgaagaagagagactggagactatgctctcccttagtctc	1258
Db	1224	ACCTGGGTGACTACATTCGCGCCGAGGAGAGATGTGGATGGATCGCCCGCTGATATTGT	1283

QY	1259	gccccgatgcagaagttataaagtcaatatttgaacaatgcccccaagcatttgt	1318
Db	1284	gcccccgacgacccaccacttccaaagaccattacttaccaaacggccccagcgcatcgcc	1343
QY	1319	aggaaagtcaaaaagatccgatattatgcaataacgaatgaaccttaagaactgtttaa	1378
Db	1344	gcgaagtacaaagagtgccgtttatgctttacaccgacgaacgttttcamaagaccggcacg	1403
QY	1379	gcatctacgaatgaatcaagaaattcttggaaccttcttatttggggaagtggaaacaca	1438
Db	1404	gcacattcacgacgacgntccggatncttggccctctctttaggggagsgtggcgacacmc	1463
QY	1439	ctgtgtattatalttaagaatcaagaacagacacatacatctaccctcaogaaatc	1498
Db	1464	ctgctgattcatctttcaagaacacagggccagacgaccttaacacattaccctccacggcatc	1523
QY	1499	actatgtccgctctctgttttcttaagaagttcccaaaagtgyltaaacattttgaagt	1558
Db	1524	accgacgtgccccctctgtacagcccgccgttccccaaagggtgaagacctgaagac	1583
QY	1559	tttccaattctgcaggagagaataatcaaatataaattggacatgactytagaaatgtgg	1618
Db	1584	ttccccacttccgtcccgccggagatcttcaatataaattggacacgctgacccgtgagagacgc	1643
QY	1619	ccaactaatcaagatctctcgtygcctgacccgctattactctagtctgttaatatggag	1678
Db	1644	cccaccagaagacgacccccctctgcttacccccgcttaccagacagcttctgtamacttgag	1703
QY	1679	agagatctagcttcaagatccatcttgccctctctcatctctctacaagaatctytagat	1738
Db	1704	cgccacctggcctccggacacgattccggccccctctctatctgttaccaaagagacgtctgac	1763
QY	1739	caaaagagaacacgaataatgtcaagaagaagaaatgcatctcttctcttatttgt	1798
Db	1764	cagcgccgacacacgattcatgtagcgacacacgacaaactatctctttcagcgtgttcac	1823
QY	1799	gagaaacgaagctgtgtaacctcacagaaatalatacaagcttctctcccaatccagcttga	1858
Db	1824	gagaaacgcgactggtatgttagaccgaamaactacacgcttctctgcccmaaccccgctggc	1883
QY	1859	gtgcagcttgaagatcccaaggttcccaagctccaacataatgacacgaatcaatgtctat	1918
Db	1884	gtgcagctggaagatcccgaggttccaaagccagacaaactatcacagacatccacggctac	1943
QY	1919	gtttttgtaagtgttggaagctgtgcaagttgttgacaggtggacatactggtacatctta	1978
Db	1944	gtgtttcagacacaccttgacagctgacacgctgtgctgtcagatgaggtgacctactggtacatctgt	2003
QY	1979	agcatltvgagcacagactgtaacctcttctgtcttctctctgtgalatalaacttcaaacac	2038
Db	2004	agcattcgccgcccagaccgacttctctgaggggtttcttccggggatnactttcaagcac	2063
QY	2039	aaaatgctctatlgaaacacacatccacctatltccatcttcaagggaatctgtctctatg	2098
Db	2064	aaagatggtgtacgagacacacctgacacctgttcccccttccgggagacgtgtgtttatcg	2123
QY	2099	tcgaatgaaaaacccaaggtctatgattctcgggtgtgcaaacatcaagacttctgggaacga	2158
Db	2124	ttctatggagaaacccccgctgttgatattcttggtgctgcacaaacagcgacttccgacacgcc	2183
QY	2159	ggcatlgacgcgcttactgaaggttctctagtctgacaaagaacactggtgatattacgag	2218
Db	2184	ggcattgaactgccctgcgaagatgtctcagctggcgaacaaacacggcgagactacttaccag	2243
QY	2219	gacagttatgaagatalttcaagcataactcttgctgagtaaaacaatgccaatggaaccaaga	2278
Db	2244	gacagctacagcgacacatctccgctactcctgctgttccaaamaacacgacatgacggcccgcc	2303
QY	2279	agcttctcccaagaaatlaaagaacctagacataggcaaaacaaatlaatgtgcaacccca	2338
Db	2304	ttcttttctccaaaactctccgccaccaccacgacgctgacaaacagatgttaacgcccacccc	2363
QY	2339	ccagttcttgaagcccatcaacagggaataactctgtactactcttccagttcaagataaag	2398







```
KEYWORDS      UNKNOWN.
SOURCE        UNKNOWN.
ORGANISM      UNCLASSIFIED.
REFERENCE     1 (bases 1 to 9009)
AUTHORS       Lollar,J.S. and Runge,M.S.
TITLE         Hybrid human/animal factor VIII
JOURNAL       Patent: US 5744446-A 1 28-APR-1998;
FEATURES      Location/Qualifiers
SOURCE        1..9009
              /organism="unknown"
BASE COUNT    2853 a 1907 c 1844 g 2405 t
ORIGIN
```

Query Match	52.9%	Score 2336.4	DB 9	Length 9009
Best Local Similarity	100.0%	Pred. NO. 0		
Matches 2337; Conservative	0	Mismatches 1	Indels 0	Gaps 0

QY	1	catgaaataagagctctccaccctgctcttcttgagctcttgagactctgctttaaagtc	60
Db	150	CATGCAAAATAGAGCTCTCACCTGCTTTCTTCTGTGCTCTTTTGCATTCCTGCTTTAGTGC	209
QY	61	caccagaataactactccctgggltgcagttggaactgtcactgagagactaatafgaaagtatct	120
Db	210	CACCAGAAGATACATACCTGGGGTGCAGTGGAACTGATGGAGCTATATGCAAAAGTGAATCT	269
QY	121	cggatgagctgctcgtlbgagcgcaagaattctcctcctagagtgccaaaatctttccatcaa	180
Db	270	CGGTGAGCTGCTGTGGACGCAAGATTTCCCTCCTAGATGCGCAAAATCTTTTCCATTCAA	329
QY	181	cacctcagtcgtgtgtaaaaaaagactctgtttgttagaatcaacggtctcaacttttaacat	240
Db	330	CACCTCATCTCTGTATCAAAAAAGACTCTTTTGTATAAATTCACGGTGTTCACCTTTTCAACAT	389
QY	241	cgtcaagccaagagccaccctgtagtgggtctgtagtctcaccatccaaaggtcgaagttta	300
Db	390	CGCTAAGCCAAGGCGCACCTTGGATGGTCTGTAGGTCTTACATTCACAGGCTGAGAGTTTA	449
QY	301	tgatacagttgttcatlaacttaagaacaatggtctccatcctgtcagttcttcaagctgt	360
Db	450	TGATACAGTGGTCATTTACATTAAAGAACATGGCTTCCATCTCTGACATCTTCAATGCTGT	509
QY	361	tgtgtatcctactcgtgnaaagctcttgaggagctgtaaatagtatgatacagaccagtc aaag	420
Db	510	TGCTGTATCTCTCACTGGAAAGCTTCTGAGAGGAGCTGAATATGATGTATGCACGACGCAAG	569
QY	421	ggagaaagaataatgataaagctctccctgggtggaaagccatacataatgctctggcagttcct	480
Db	570	GGAGAAAGAAATGATATAAGTCTTCCCTGGGGAAGCCATATGCTCTGGCAGGCTCT	629
QY	481	gaaagaagaatgtccaatgycctctgaccaactgyccttcaactcactatatacttctcca	540
Db	630	GAAAGAGAAATGCTCAATGGCTCTTGACCCACATGTGCTTACTACTCATATCTTTTCTCA	689
QY	541	tgttgagacctgtgtaaaagacttgaaltcaagagctcaatltbgagccctactagtaatgtagaga	600
Db	690	TGTGAGACTGTGTTAAAGACTTGAATTCAGGGCTCATTTGGAGGCCCTACATGATATGTAAGA	749
QY	601	agggaagctctggccaagaaaagaacagaaactgtgcaaaattatatactacttltgtctgt	660
Db	750	AGGGAGCTCTGGCCAAAGGAAAGACACACACACTTGCACAAATTTATACATCTTTTGTGCTGT	809
QY	661	atttgatgaagaaggaaaatttgagcctcgaaaaacaaagaagactccctgtatggcagataagga	720
Db	810	ATTTGATGATGAAGGAAAATTTGGCACTTCGAAACAAAGAACTCTTGTATGCAAGGATAGGGA	869
QY	721	tgtctgacatctctcgggctgtgacctaaatatgacaacagtc aaatgtgtatgttaaacaggtc	780
Db	870	TGCTGCAATCTCTCTGGGCTTGCCCTTAAATATGCACACATCATGTGTATGTAAACAGGTC	929
QY	781	tctgcacagttctgatttgatgccacaagaaaatcagttcatattgtgcatgtgatttgatgg	840

Dd	930	TCGCGCAGGCTGATTGGAGTCCACAGCAAAATCAGTCCTATTGGCATGTGATTGGAATGGG	989
Qy	841	caccacctcttgaagtgcaactcaatattccctcgaaggtcaacaattctctgttgaagaca	900
Dd	990	CACACACTCTGAAGTGCACCTCAATATTCTCGAAGGTCAACACTTCTGTGTAGGACACA	1049
Qy	901	tcgcgaagcgctccttggaaatctgcgaataacttctcctaactgtctcaacactcttgat	960
Dd	1050	TCGCGCAGCGTCTCTGGAAATCTCGCCAAATACCTTCCCTACTGCTCAACACTCTGTAT	1109
Qy	961	ggaccttggacaagtttctaactgtttgttaatactcttcacccaacatgatgaaatga	1020
Dd	1110	GGACCTTGAGACAGTTTCTACTGTTTGTCAATCTCTCCACACAACAGATGGCATGGA	1169
Qy	1021	agctatctcaaaatagaacagctgtccagaaggagaccacaactacgaatgaataaatga	1080
Dd	1170	AGCTTATGTCAAAGTACACAGCTGTCCAGAGGAACCCCAACTACGATATAAAATATATGA	1220
Qy	1081	agaagcggaaagactatgatgatcttaactgaattctgaatltgaatltgttcaagtttga	1140
Dd	1230	AGAAGCGGAAGACTATGATGTGATCTTACTGATTCGTGAAGATGATGTGTGACGGTTTGA	1280
Qy	1141	tgaatgaacaactctcctcctcttataccaactctgcctaagtgtgccagaagcaactcaaac	1200
Dd	1290	TGATGACAAACTCTCTCTTATCCAAATTCGCTCAAGTGGCCAGGAAACATCCCTAATAAC	1349
Qy	1201	tctgggtacaataactgtctctctgaagaaggagactgtggacttgcctccctagtctctgc	1260
Dd	1350	TTGGGTACATTACATTCCTCTCTGAAAGGGAGACTGGACATAGCTCCCTTATGCTCTCGC	1409
Qy	1261	ccccatgaacgaagattataaaagttcaataattgaacaatlygcccccaacggaattgtatg	1320
Dd	1410	CCCCGATGACGAAGATTATAAAGTCAATTTTGAACAAATGAGGCCCTCACGGGATTTGTATG	1460
Qy	1321	gaagtcacaaaaaagtcggatttalygcatacacagaatgaacaccttaagaactcgttgaagc	1380
Dd	1470	GAAGTACAAAAAAGTCCGATTATAGGATATCACAGATGAACCTTTTAAAGATCTGTGAAGC	1520
Qy	1381	tattcagatgaataacagaatctctggaccttaacttttbbgggaagtttggagaacact	1440
Dd	1530	TATTTCAGATGAATACAGGAATCTTGGACCTTTACTTTATGGGGAATGTGGAGACACT	1580
Qy	1441	gttattatatttlaagaatacaagcaagcagaacataatacaattaccctcacaaggaataac	1500
Dd	1590	GTTGATTATATTTAAGATACACAGCAAGCACAACATATACATCTACCTCACAGGAATCAC	1649
Qy	1501	tgatgtccgtcctcttgtatccaaggagattaccaaaagltgaanaacatttgaagattt	1560
Dd	1650	TGATGTCCGTCCTTTGTATTCAAGAGGATTTACCAAAAGGTGTAAACATTTTGAAGGATTT	1709
Qy	1561	tccaattctgcaggagaataattcaaatataatgaacagtgactgttgaagaatlyggcc	1620
Dd	1710	TCCAATTCTGCGACAGAAATATTTCAAATTAATGTGAAGAGTGTGAAGATGGGCC	1769
Qy	1621	aaacaaatcaaatctcgtgtgacctgaacccgctaactactagtlttcgttaaatlyggaag	1680
Dd	1770	AACATAATCAATCTCTGGTGCTGACCGCGCTATTACTAGTTTCGTTAATATGTGAAG	1820
Qy	1661	agatctagcttcaggactcatlyggccctctcctcatctctgtaacaaagaactctgtatgata	1740
Dd	1830	AGATCTTAGCTTCAGAGCTCATTTGGCCCTCTCCATCTCTCAAAAGAAATCTGTGTAGATCA	1889
Qy	1741	aagagggaacacagataatgtcagaacaaagggaatgtcatctcgttlttctgtatlttgatga	1800
Dd	1890	AAGAGGAACCCAGTATATGTGCACAGAGGAATGTCACTCGTTTTCGTATTGTGAATGA	1949
Qy	1801	gaacccgaagctgtgaccttcaagagaataataaacgttctccccaatccaaactcggaggt	1860
Dd	1950	GAACCGGAAGCTGTGACTCTACAGAGATATACAGCCTTTCTCCCAATCCAGCTGGAGT	2009
Qy	1861	gcaagcttgaagatccagaagtttccaagcctccaacatcatatgcaacagatccaatlygtatgt	1920
Dd	2010	GCACCTTGAGAGATCAAGATGTTCCAAAGCTGTCAACATCATGCAAGCATCAATATGGCTATGT	2069



QY	1921	ttcttgtaagtcttgcagttgtcagtttcttgatgagagtggtgacatactcgtgatacttctaa	1980
Db	2070	TTTTTGATAGATTTCAGATTGCAGTTGCTTTGTCATGAGAGGTGGCAATCGTTCATTTCAAG	2129
QY	1981	catgtagagccagacagctgacttcctctctctctctctctctctctctctctctctcaaacaa	2040
Db	2130	CATTGGAGCACAACACTGACTTCCTTTCTTGCTTTCTTCTTGATATACCTTCAAAACACA	2189
QY	2041	aatgtctctatgaaagaaacactccacccatccatcccatctcagagagaaactgtcttcattgc	2100
Db	2190	AATGCTCTATGAGAGACACTCAACCCATTGCCATTTCCATTCCTGAGGAGAAACTGCTTCATGTC	2249
QY	2101	gatgtagaaacccagctctatgattcttgaggtgtccacaactcagaacttcggaaacagaag	2160
Db	2250	GATGGAAAAACCCAGGCTCTATGGAATTCTGGGGGTGGCCAAACACTCAACATTTCCGAAACAGAGG	2309
QY	2161	catgagccgctctactgaaggtctcttaagtcttgaaagaagaaacgtgtgattatctacagga	2220
Db	2310	CATGACCGCCTTACTGAGGTTTCTTAAGTTTGATGCAAAACAACCTGTGTGATTATACGAGGA	2369
QY	2221	caagctatgaaagatatctcagcatactctgcgtgagtaaaacaatgcattgacacaagaag	2280
Db	2370	CAGTTATGAAGATATTTCACACATACCTTGCGTAGTAATAAAACAATGCCAATGTGAACCAAGAG	2429
QY	2281	ctctctccgaattcaagacacccctagcactaggcaaaagcaatttatgcaccacca	2338
Db	2430	CTTCTCCCAGAATTTCAGACACCCCTTGACCTTAGGCAAAAGCAATTTATGCCACACAC	2487

Db	390	CGCTAAGCCAAGGCCACCCTTGATGGGTCTGCTAGGTCCTACCATCCAGGCTGAGGTTA	443
Qy	301	tgatacaagtggkclataacttaagaacatgagcttccacatcgtltaagcttaacatgclgt	360
Db	450	TGATCACTGGTCATTTACACTTAGAACAATGGGCTTCCCATCTGCTAGTCTTCACTGCTGT	509
Qy	361	ttgtgtlatacctaacttggaaagcttcgtgaggagcttgaaatgatatgatatcaagacagttcaag	420
Db	510	TGGTGTATTCCTACTGGAAAGCTTGTGAGGAGCTGATATGATGATGATCAACAGCAGTCAAG	569
Qy	421	ggagagaagaagaatgaataatcttcccttggtgagaagcatalatatgtcttggcagtcct	480
Db	570	GGAGAAAGAAGATGATATAAGTCTTCCCTGGTGGAAGCATACATATGCTGTGGCAGAGTCTT	629
Qy	481	gaagaagaatgttcaatgtgctcttggcccaacgttgcttactatactatactcttcca	540
Db	630	GAAAGAAGATGGTCCATAGGCGCTTGACCCACGTGGCTTACTCATCTCAATCTTTGCA	689
Qy	541	tgtggaacttgtlaaagaacttgaataaagccatctgaagccatctagtagttglaagga	600
Db	690	TGTGACCTTGTTAAAGACTTGAATTTAGGCGCTCATTTGGAGCCCTACTGATGTATGAGA	749
Qy	601	agggaagcttggccaaggaagaacacagacacttgcacaattatatactaacttttgcgt	660
Db	750	AGGAGTGTGGCCCAAGGAAAGACACGACACTTGCACAAATTTATATCTTTTGGCTGT	809
Qy	661	atttgatgaaggaagaagtgtgcaactcagaacaaagaactccttgatgcaagatlaagga	720

	RESULT	12				29-SEP-1999
	AR029065					
	LOCUS	AR029065	9009 bp	DNA	PAT	
	DEFINITION	Sequence 1 from patent US 5859204.				
	ACCESSION	AR029065				
	VERSION	AR029065.1	GI:5941038			
	KEYWORDS	.				
	SOURCE	Unknown.				
	ORGANISM	Unclassified.				
	REFERENCE	1 (bases 1 to 9009)				
	AUTHORS	Lollar,J.S.				
	TITLE	Modified factor VIII Patent: US 5859204-A 1 12-JAN-1999;				
	JOURNAL	Location/Qualifiers				
	FEATURES	1..9009				
	SOURCE	/organism="unknown"				
	BASE COUNT	2853 a	1907 c	1844 g	2405 t	
	ORIGIN					

QY	721	tgcgcagctcgcctcggagcgcgcgtacaaatgcacacagtcgatgttatgtaacaagtc	780
Db	870	TGCTGCATCTCTCTGGGCGCTGGCTTAAAAAGCACACAGTCAAAGGTATGTAACAGGTC	929
QY	781	tctgcacagctctgattcgatgcacacgaagaaatcagtcctatctgcacgtaatgcgaatgcg	840
Db	930	TCTGCACAGTCTTGATTGGATGCGCCACAGGAATTCAGTCTATTGGCAGTATTTGGAAATGG	989
QY	841	caccacctccgcgaagtcgaactcaatatctctcgaaagtcacacaaatctcttgcggaacca	900
Db	990	CACCACTCTCAAGTCGACCTCAATATTCCTCGAAAGGTCACACTTTCTGTGAGGACCA	1049
QY	901	tgcgcagcgcgcctctgcgaaatctgcgcaaatctcttcctactgcgtcgaacactcttgat	960
Db	1050	TGCGCAGCGCTCTTTGGAAATCTGCCAATAACTTTCTTACTGCTCAAAACCTTTGAT	1109
QY	961	ggaccttcggacagttcttactggttttgcatalactctccacacacacgatgcgcatgta	1020
Db	1110	GGACCTTGGACAGTTTCTACAGTTTTCATATCTCTCCACACACATGATGGCATGGA	1169

Query Match	Similarity	52.9%	Score 2336.4	DB 9	Length 9009
Best Local	Similarity	100.0%	Pred. No. 0		
Matches 2337	Conservative	0	Mismatches	1	Indels 0
					Gaps
Qy	1	catgcaaatagagctctcacctgctcttcctcgtgccttttgagcttcagcttagtc	60		
Db	150	CATGCAAAATAGAGCTCTCTCACCTGCTTCTTTCTGTGCTTTTGGGATTCGCTTTAAGTC	2099		
Qy	61	caccagaagaatacctactcctgggtgcagatggaactgcataggagactatatatgcaaaagtatct	120		
Db	210	CACCGAAGATATCTACTCTGGGTGCGAGTGCACTGCAATGGAGACTATATGCAAAAGTATCT	269		
Qy	121	cgtgtagctgcctctgtgagcgcgaagattctccctccctagagtgccaaatctttccatctcaa	180		
Db	270	CGGTAGCTGCTGTGGAGCGCAAGTTTCCCTCAAGATGTGCCAAAAAATCTTTTCCATTGCA	329		
Qy	181	caaccctagctgtgatacaaaaagacctctgtttgtagaattcaacgtttacacctttcaacat	240		
Db	330	CACCTCAAGTCTGTTCAAAAAAGACTCTGTTTGTGTGAATTCACGGTTACCTTTTCAACAT	389		
Qy	241	cgttaagccaagcgcaaccttgatgtgctctgcttagtgcctacccatccagagctgaagtta	300		

Qy	1141	tgatgacaactctccctcccttaccacaattcgctcagttgcccagaagcaltccataaac	1200
Db	1290	TGATGACACACTCTCTTCCTTTATCCAAATTCCTCAGTTGGCCAAAGAACCTCTTAAC	1349
Qy	1201	ttgggttaactaaccttgcctctgaagaaggagacttggactatgtctcccttaagtcctgc	1260
Db	1350	TTGGGTAAATTACATTCCTCTCTGAAAGGGAGACTGGGACTATMGCTCCCTTAGTCCTGC	1409
Qy	1261	ccccatgacgaagttatataagtcacatttggacaatggccctcagcagattgtag	1320
Db	1410	CCCCGATACAGAAAGTTATTAACATCAATTTTGAACAAATGGCCCTCAAGCGATTGGTAG	1469
Qy	1321	gaagbacaataaagctccgattttagtcatacacagatgaacaccttaagactctgtaagc	1380
Db	1470	GAAGTACAAAATAAGCTCCGATTATAGGATTCACAGATGAACCTTTAAGACTCGTAGAGC	1529



QY 1381 taccagatcaatcaagaatcttggacccttaacttaattgagggaagtggagacact 1440  
|||||  
Db 1530 TATTACAGATCAATCAGAGAACTTGGACCTTTATGGAGAGTTGGAGACACACT 1589  
QY 1441 gtgatataatlaagaatcaagaacagacacataaacttaacctcaacgaatcac 1500  
1590 GTTGAATATATTTAAGATCAAGACAGACATATACATCTACCTCAGGATCAC 1649  
QY 1501 tgatgtccgtcttctgtaatacaagaagattacaaaagtgtaaaacaattgaagattc 1560  
1550 TGATGTCCGTCCTTTGATTCAGAGATTAACAAAAGTGCTAAACATTTGAAGGATTT 1709  
QY 1561 tccaatctgcagagagaatataatcaataataatgaagtgagctgtagaagaatggcc 1620  
1710 TCCAAATTCCTGCAGAGAAATATTCAAAATTAATTAATGACATGCTAGTAGAGATGGCC 1769  
QY 1621 aactaaatcaagatcctcgtgctcgaaccgctataacttaactgattcgtlaataatgagag 1680  
1770 AACTAAATCAAGATCTCGTGCTGACCCGCTATTACTGATGTTGCTTAATATGAGAG 1829  
QY 1681 agatctagctcaagaactcattggccctctccatctgtaataagaatctgtatga 1740  
1830 AGATCTAGCTTCAGAGACTCATTTGGCCCTCTCCATCTGCTACAAAGATCTGTAGATCA 1889  
QY 1741 aagaagaaacagataatgtaagaagaagaagatgtaacatcctgttctctgattatga 1800  
1890 AAGAGAAACAGATTAATGTCAGACAAAGAGATGTCATCTGTTCTGTATTGTATGA 1949  
QY 1801 gaacgaagctgtaacctcaagaataatacaagcttctcccaatccagcttgagat 1860  
1950 GAACCGAAGCTGGTACCTTCACAGAGAAATATACAGCGTTCTCCCAATCCAGCTGGAGT 2009  
QY 1861 gcagcttgagagatccagaagttccaagctccaacatcatgcaagcatcaatggtatgt 1920  
2010 GCACCTTGAGAGATCAAGATGTCAGAGCTCCACATCATGACAGACATCAATGCTATGT 2069  
QY 1921 ttgtgagattgagctgtaagttgttgacagagtgagctgtaacatgtaacatc 1980  
2070 TTTTGATAGTTTGCAGATGTCAGTTGTTTGTGATGAGTGAGTGATCATTTCTTAAG 2129  
QY 1981 cattgagacagacagactccttctgcttcttcttcttctgtaataacttaacacaa 2040  
2130 CATGAGACACAGACTGACTCTCTTCTGTCCTTCTCTGATATATACCTTCAACAA 2189  
QY 2041 aatgcttataagaagacacactcaacctatccatctctcagagaaactgtctcattgtc 2100  
2190 AATGCTATGAAGACACACTCACCTATTCCTCATCTCAAGAGAAACTGTCTCATGTCT 2249  
QY 2101 gatgaaacacacaggtctatgagattctcgggtgacacacacagacttctcggaaacagag 2160  
2250 GATGAAACACCCAGGTCTATGAGATTCTGGGTTGCCACACTCAACATTTGCGAAGAGG 2309  
QY 2161 catgaccgcttactgaagtttctagttgtgacaagaacactgtaattatlaacagaga 2220  
2310 CATGACCGCTTACTGAAGTTTCTAGTTGTGACAAGAACTGCTGATTAATTCAGAGA 2369  
QY 2221 caattataagaagataattcagacatactgtcgtgataaaacaatgcaattgaacaaag 2280  
2370 CAGTTATGAAGATATTTACACATATCTGCTGAGTAAAGAAATGCTGATGAACCAAGAG 2429  
QY 2281 ctctcccaagaatcaagaacacctagcaactaaggcaaaagaatlaattccaccca 2338  
2430 CTCTCCCAAGATTCAGACACCTTAGCACTAGGCAAAAGCAATTTAATGCCACACACA 2487  
Db

RESULT 13  
LOCUS I31901 9009 bp DNA PAT 06-FEB-1997  
DEFINITION Sequence 3 from patent US 5583209.  
ACCESSION I31901  
VERSION I31901.1 GI:1822692  
KEYWORDS

SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 9009)  
AUTHORS Lollar,J.S. and Runge,M.S.  
TITLE Hybrid human/porcine factor VII  
JOURNAL Patent: US 5583209-A 3 10-DEC-1996;  
FEATURES  
source location/qualifiers  
1..9009 /organism="unknown"  
BASE COUNT 2853 a 1907 c 1844 g 2405 t  
ORIGIN  
Query Match 52.9%; Score 2336.4; DB 10; Length 9009;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2337; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 catgaaatagagctctccacatgcttcttctcgtgcttctgagattcgttaagtgc 60  
|||  
Db 150 CATCAATTAAGACTCTCCACTGCTTTCTGTGCTTGTGCGATTCTGCTTAGTGC 209  
QY 61 caccagaagatactacctggtgagtggaactgtaactggaactatataatgcaagtgtact 120  
|||  
Db 210 CACCAGAAGATTAACCTGGGTGCAGTGAACCTGATGGAGATATATGAAAGATCT 269  
QY 121 cgtgtgagctgctgtgagcgaagattctcctcctagagtgccaaatcttccatcaa 180  
|||  
Db 270 CGGTGAGCTGCTGTGACGCAAGATTTCTCCTAGAGTGCCTCAAAATCTTTTCAATTCAA 329  
QY 181 caactagctgtgtaacaaagaactcgtgtgtgagaaatcaagttcaacttccaacat 240  
|||  
Db 330 CACTCAGTCTGTACAAAAGACTCTGTTGTGAATTCACGGTTCACCTTTTCAACAT 389  
QY 241 cgttaagcacaagcgcacctggaatggtctgctagttcttaacacacagcgtgagttta 300  
|||  
Db 390 CGTTAAGCACAAGGCGACCTGGATGGGTCTGTAGTCTACATCCAGGCTGAGGTTTA 449  
QY 301 tgataagtgtaatttaacttaagaacatggtctccatctcctgtcagttctcattgtc 360  
|||  
Db 450 TGATTAAGTGTGATTAACATTAAGAACATGCTTCCATCTCTCACTTTCATATGCTGT 509  
QY 361 ttgtgttccctcaaggaaagcttctcggagggagcgaatatatgaacagacacaaag 420  
|||  
Db 510 TGTGCTATCTACTGGAAGCTTCTGAGGAGCTGGAATATATATATGATCAGACAGTCAAG 569  
QY 421 ggaagaagaagaatgaatcttccctggtggaagccatacatatgctcgtgagttcct 480  
|||  
Db 570 GGAGAAAGAAAGATGATTAAGCTTCCCGTGGGAAGCCATATATGTCGCAAGTCT 629  
QY 481 gaaagaagaatggtccaatggtccttgacccacatgctcttaactcaatatcttctca 540  
|||  
Db 630 GAAAGAAATGTCATATGCTTGAACCTGACCCACTGTGCTTACATATATCTTTTCA 689  
QY 541 ttgtgacctgtgtaagaactgaaatccaggtcctatcttgagcccttaactagatgtaga 600  
|||  
Db 690 TGTGACCTGCTTAAAGACTTGAATTCAGGCTCTCATTTGAGACCTACTAGATATGAGGA 749  
QY 601 agggagcttggtcacaaggaagaacacagaccttgacaaatlaactactcttctgtct 660  
|||  
Db 750 AGGAGTCTGGCCAAAGAAAGACACAGACTTGCACAAATTTATACACTTTTGGCTGT 809  
QY 661 atttgatgaagggaaagtgtgcaactcaagaacaaagaactccttgatgcaagatagga 720  
|||  
Db 810 ATTTGATGAAGGAAAGGTTGGCACTCAGAAACAAAGAACTCTTGATGCGAGATAGGA 869  
QY 721 tgcgtcatctgctcgtggccttgctaaatgcaacagctcaatgtaatgtaaaacagctc 780  
|||  
Db 870 TGTGCTCATCTGCTGGGCTGTGCTTAATATGACACAGTCAATGTTATGTAACAGTCT 929  
QY 781 tctgccaagttcgtatgtaatgcaagaagaatcagctcattgtgcatgtaattggaatgg 840  
|||  
Db 930 TCTGCCAGGCTGTGATGATGCCACAGGAATCAGTCAATGTCATGTGATTTGAATGGG 989



QY	841	caccactctgtgaagtgcactcoactaatactccctcgaaagtgcaacatllctltgtgaaggaacca	900
Db	990	CACCACACTCCTGAAGTGTGCACTCAAAATATTCCTTCGAAGGTCACACATTTCTTGTGAGGAACCA	1049
QY	901	tgcacagagctcccttgnaaatctgcacaatbaacttctctactctgtctaaanaactctgat	960
Db	1050	TCGCCACAGGCTCCTCTGGAAATCTGCGCAAAATACATTTCTCTACTGCTCAACACATCTTAT	1109
QY	961	ggaactctggaacgtlttctactgtlttgcatatacttctcccaacaacatga tggcatga	1020
Db	1110	GGACCTTGAGACAGTTTCTACTGTTTGTGCATATCTCTTCCACCAACATGATGGCATGGA	1169
QY	1021	agcttatgttcaagtagaacagctgtgccagagaaaccccaactacgaatgtaaataatga	1080
Db	1170	AGCTTATGTCAAAAGTAAACAGCTGTCCAGAGAAACCCCACTACGAAATGAAGAAATTAAGA	1229
QY	1081	agaagcggaaagactatgatgatga tcttactgattcttgaatgtgatgtggtcagaatttga	1140
Db	1230	AGAGCGGAAGACTATGATGATGATCTTACTGATTTCTGAATGSGATGTGGTCAGGTTTGA	1288
QY	1141	tga tgaacaactctctctctcttcaacaaatgcgtcagltgtgccaaagaacatccaaaac	1200
Db	1290	TGATGACAACTCTCTCTCTTATTCCAAATTCGCTCAGTTGCCAAGAAAGCATCTTAAAC	1349
QY	1201	ttaggtatcaattaaacttgctgtctgaagagggaatggagactatgctcccttaagctctgc	1260
Db	1350	TTGGGTACATTAATATGCTCTCTGAAGGAGAGACTGGGACTATGCTCTTATGTCTGC	1409
QY	1261	ccccgatgacaagaagtataaaatcaaatltgaacaatgacctcaagcgatgtgtgag	1320
Db	1410	CCCCGATGACAGAGATTATTAAGAACTCATATTTAAACATATGGCCCTCAGCGGATTGGAG	1469
QY	1321	gaagtaacaaaaagctccgaattatlygcatacacagatgaaaccttlaagactcgtlgaagc	1380
Db	1470	GAAGTACAAAAAAGTCGCGATTTATGGCATACACAGATGAACCTTTAAGACTCGTGAAGC	1529
QY	1381	tatcagcatgaatcaggaactcttgggaacttcaactttagtgggaagtltgaaacaact	1440
Db	1530	TATTTGACGCAATGATCAAGAAATCTTGGGACCTTCTTATGSGGAAGTTGGAGACACT	1589
QY	1441	gttgatatataattaaagatacaagcaagcagacaataaatacctcaacccacggaaatcac	1500
Db	1590	GTTGATTATATTTTAAGATATCAAGCAACACAGCCATATACACTTCACTCACGGAATCAC	1649
QY	1501	tga tgcgcgtcccttgytatccaagggagatiaccaaaagtgltaaacaatttgaagatct	1560
Db	1650	TGATGTCCGCTCTTGTGATTTCAAGGAATTAACCAAAAGGTGTAACATTTGAAGGATTT	1709
QY	1561	tccaattctgcagggagaaatattcaaatataa tlgaaacagtgactgttagaagaatlgagcc	1620
Db	1710	TCCAATTCGCGCAGGAATAATTTCAATATAATTAATGAACGTGACTGTAGAGAAATGGGCC	1769
QY	1621	aactaaataagaa tccctgcgtgcctgacccgcatacttactctagttctgtlaa tlgagag	1680
Db	1770	AACTAATGATGATCCTCGGTGGCGTGAACCCGCTATTACTGTAGTTTCGTTAATGTGAGAG	1829
QY	1681	agatctagcttcaaggaactatggccctctccatctctgtactcaagaagaatctgtatagatca	1740
Db	1830	AGATCTAGGTTCTAGGAGCTCATTTGGCCCTCTTCCTCATCTGTACACAAAGATCTGTAGATCA	1889
QY	1741	aagagaaacccaaataatgtcacaacaaaggaagatgatactcgttcttctgattatgtga	1800
Db	1890	AAGAGGAACCAATATATGTCACACAAAGAGGAATGTATCTCTGTATTGATGATGA	1949
QY	1801	gaaccgaagctgtgtactcacaagaaatacacaagcttctctcccaactcagcttgagt	1860
Db	1950	GAAACGAACCTGGTACTCTACACAGATATATACAAAGCTTTCTGCCCAATCAGCTGAGGT	2009
QY	1861	gcagcttggagatcccaaggttccaagcctccaacatcatgcaacagcatlcaatlygtctat	1920
Db	2010	GCAGCTTGGAGATCCAGAGTTCCAAGCCTCAACATCATGATCAGCAGCATATGCTATGT	2069

OY	1921	tlttgaatgttgagctgctgcagtgttgtttgttgatbagygfgagataactcgtlaattctaag	1980
Db	2070	TTTTGATATTGGACAGTTCTTCAGTGTGTTGGATGAGTGCGATCTCGTACATTCTTAAG	2129
OY	1981	cattgagagaacagactcccttcctctcgtctctctctctctgtatataaccttaaacacaa	2040
Db	2130	CATTGAGACACAGACTGACTCTCTTCTTCTGTCTTCTTCTGTGATATFACCTTCAACACAA	2189
OY	2041	aatgtgtctatgaagacacacitcaccttatcccatctctcaaggaaaactgttcttcagtc	2100
Db	2190	AATGGCTATTGAAGACACACTCACCCCTATTGCCATTCTCAGAGAAGAACGTCCTCATGTGC	2249
OY	2101	gatbgaaaaaccagtgctctatlgatctctggggtggtcgccaacttaagctcttcgaaacaag	2160
Db	2250	GATGGAAAAACCAGGCTATTGGAATTTCTGGGGGCCACAACCTCAGACTTTGGAAACAGAGG	2309
OY	2161	cattgaccgcttactbaaaggttcttaagttctgtgtgacaagaacactggtgtattatcagaga	2220
Db	2310	CATTGACCGCTTACTATAAGTTCCTTAGTTGTGCACAAGAACACTGGTGATATTACGAGAGA	2369
OY	2221	cagttatgaagaatlattcagacatacttgcgtsgttaaaaaaalgtccattgaaccaagaag	2280
Db	2370	CAGTTATGAAGATATTTCAGCATACTTCTGTGATTAATAAACAATGCCATTGAACCAAGAAG	2429
OY	2281	cttctcccgaattcaaagcaccccttagcactctggcgaaaaagcaattaatgccaccaca	2338
Db	2430	CTTCTCCCAGAAATTCCAAGACACCCTTAGCACTAGGCAAAAAGCAATTTAATGCCACACACA	2487
RESULT	14		
LOCUS	163424	9009 bp	DNA
DEFINITION	Sequence 1 from patent US 5663060.	PAT	07-OCT-1997
ACCESSION	163424		
VERSION	163424.1	GI:2480997	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	Unclassified.		
AUTHORS	1 (bases 1 to 9009)		
TITLE	Lollar,J.S. and Runge,M.S.		
JOURNAL	Hybrid human/animal factor VIII		
FEATURES	Patent: US 5663060-A 1 02-SEP-1997;		
	Location/Qualifiers		
	1..9009		
	/organism="unknown"		
BASE COUNT	2853 a	1907 c	1844 g
ORIGIN		2405 t	
Query Match	52.9%; Score 2336.4; DB 10; Length 9009;		
Best Local Similarity	100.0%; Pred. No. 0;		
Matches 2337; Conservative	0; Mismatches 1; Indels 0; Gaps 0;		
OY	1	catlgaataaagagctctccactctgtcttcttctgtgcctllttggatctctgtatagc	60
Db	150	CATGCAAAATAGAGCTCTCCACCTGGCTTCTTCTGTGCTTTTGCAATTCGCTTAAAGTC	209
OY	61	caccgaagaatacctaccbtggfycagbtgnaactgfcatalggagaciatatgcaagfyact	120
Db	210	CACCGAAGATATCTACCTGGGTGTCAGTGGAACTGTCATGAGACTATTATGCAAAAGTATCT	269
OY	121	cggtagagctgcgcctgtgagcgcagaatatctccctcctlagagtgccaanaatcttlccalc	180
Db	270	CGGTAGAGCTCGCTGTGGAGCGCAAGATTTCCTCTAGAGTGGCCAAATCTTTCCATTCAA	329
OY	181	caactcagtcgtgtacaataaagaactcgttlttgtagaatlcacgyltcaoctlltacaat	240
Db	330	CACCTCAGTCGTGTACAAAAGACACTCTGTTGTAAATTAACGAGTTTCAACAT	389
OY	241	cgttaagccaagccacccctgtagtgggtctctctagtgctctacacccaagctgaggtlla	300
Db	390	CGCTAAGCCAAAGCACCCCTGTGATGGGTCTGTGTAGTCTCTACATCCAGGCTGAGCTTTA	449



Qy	301	tgtataagtggtcaataacttaacctaagaacatgctcccaactccctgtgcaacttcbaatgctgt	360
Db	450	TGATACAGTGTCTATTACACTTAAAGACATGGCTTCCCTGCTGCAAGTCTTCAAGTGT	509
Qy	361	tgtgtatctctacttgaagaagctcttgaaggagctgaatatatgtatacagacagtcacaag	420
Db	510	TGtGTATCCCTACTGGAAACCTTCTGAGGAGAGTGATATGATGATCAGACAGTCAAG	569
Qy	421	ggagaaagaagatgataaagtcctctccctgtgtggaagccaataatgtctcgtgcagttcct	480
Db	570	GGAGAAAAGAAAGATATAAATCTTCCCTGGTGGAAACCATATCATATGTCTGGCAGGTCT	629
Qy	481	gaagagaatgtgtccaaatggccctctgacccaactgtgcccctactcctatatctcttcca	540
Db	630	GAAGAAGAAATGTGCCAATGSCCTCTACACCCTGTGCTACTCTATATCTTTCTTCCA	689
Qy	541	tgtgagacctgtgtataaagaacttgaatcagaagctcatgtgaagccctactagatgtlaaga	600
Db	690	TGTGTGACCTGTGTAAAGACTTGAAATTCAGAGCCCTCATTTGAGGCCCTACTAGTATGAGA	749
Qy	601	agggaagctgtggccaaggaaagaacaaagacctgtgaccaaattataactacttltgtcgt	660
Db	750	AGGGAAGTCTGGCCAAAGAAAAGACACAGACTTGCAAAATTTATATCATCTTTTGTCTGT	809
Qy	661	atttgaagaagggaaaaagtgtgcaaccacagaaaaagaacatccctgtatgacagataagga	720
Db	810	ATTTGTATGAAGGGGAAAAGTGGGCATCTCAGAAAACAAAGAACTCTTGATGACAGATAGGA	869
Qy	721	tgtctgcatctgtctcgagctgtgcctataaatgcacacagtcacatgtgtatgttaacaagtc	780
Db	870	TGCTGCATCTGCTCGGGCCCTGGCTTAAATGACACAGTCAATGGTTATGTAAACAGGTC	929
Qy	781	tctctcagagctctatgtgatatccacaggaatacaatcctatctgacatgtgatatgtgaatgg	840
Db	930	TCTGTCCAGTCTCATTTGAGATGCCACAGAAATCAGTCTATTTGGCATGTGATTTGGAAATGG	989
Qy	841	caaccactctgaagtgcaatcacaatatctctcgaaggtcaacacatcttctgtgagaacca	900
Db	990	CACCACTCTGGAAGTGCACACTAATTTCTCCAGAGTCCACACATTTCTGTGAGAGACCA	1049
Qy	901	tcgcagagagctcttggaaatccgcgcaataactctctactatgtctgaacaaccttgaat	960
Db	1050	TCGCGACAGCCTCTTGGAAATCTCGCCAAATCTTTCTTACTGCTCAACAACTCTTGAT	1109
Qy	961	ggaaccttgacagtlctctactgttltgtlcatatactcttccaccaacaatgaatgtgcaltga	1020
Db	1110	GGACCTTGGACAGTTTCTACTGTGTTTGTCAATNTCTTCCCAACATGATGGGATGGA	1169
Qy	1021	agcttatgtctaaagttagacagctgtctccagaagaaccccaactagaatgtaaaataatga	1080
Db	1170	AGCTTATGTCTAAAGTGAACGCTGTCCAAAGGAAACCCCAACTACCAATGAAGAAAATAAGA	1229
Qy	1081	agaagcggaagactatgatgatgcatcttactatctcgtgaatgagatgtgtcgaagtgtga	1140
Db	1230	AGAAGCGGAAGACTATGATGATGATGATCTTACTGATTTCTGAATGAATGATGTGGTCAAGTTGA	1289
Qy	1141	tgatgatcaactctccttcccttataccaatctgcgtcgaagttgtgccaagaagacacctaaac	1200
Db	1290	TGATGACAACTCTCTCTCTTATCTCCAAATTTGGCTCAGTTGGCAAGAAGCATCTCTAAAC	1349
Qy	1201	tttggtatcatctaatctgtcgtctbaaagaagaagacatggaagactatgtctcctaagtcctgcg	1260
Db	1350	TTGGGTATCATTAATTTGCTGCTGAAGAGAGAGACTGGGACTATGTCTCCCTTAGTCTCTGC	1409
Qy	1261	ccccgatgacagaagatataaagaatcaatatatttgaacaatgtgcccctcagcgaattgttag	1320
Db	1410	CCCCGATGACAGAAAGTTATAAAGTCAATATTTGAAACAATGGCCCTCAGCGAATGGTAG	1469
Qy	1321	gaatatcaaaaaagtcogattttttgtgtcatcacagatgaatgaacactttaagactcgtgaagc	1380
Db	1470	GAATATCAAAAAAGTCGGAATTTATGGCATACAGTGAACCTTTAAGACTCGTGAAGC	1529

[illegible]

RESULT	15
HOMFEVIII	
LOCUS	HOMFEVIII 9009 bp mRNA PRI 08-NOV-1994
DEFINITION	Human coagulation factor VIII:C (anti-hemophilic factor) mRNA.
ACCESSION	K01740
VERSION	K01740.1 GI:182802
KEYWORDS	factor VIII; hemophilia.
SOURCE	Human cDNA to liver mRNA, and genomic DNA.



ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 9009)
AUTHORS	Toole,J.J., Knopf,J.L., Mooney,J.M., Sultzman,L.A., Buecker,J.L., Pittman,D.D., Kaufman,R.J., Brown,E., Shoemaker,C., Orr,E.C., Amphlett,G.W., Foster,W.B., Coe,M.L., Knutson,G.J., Fass,D.N. and Hewick,R.M.
TITLE	Molecular cloning of a cDNA encoding human antihaemophilic factor
JOURNAL	Nature 312 (1984)
COMMENT	Currently hemophilics are treated with factor VIII:C purified from blood. The human factor VIII:C cDNA has now been isolated, cloned and expressed. The gene was also isolated. It consists of a large number of exons spanning over 180 kb. The authors kindly supplied the sequence directly to Genbank by computer network.
FEATURES	Location/Qualifiers
SOURCE	1..9009
gene	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/map="Xq28"
	151..7206
	/gene="F8C"
CDS	151..7206
	/gene="F8C"
	/note="factor VIII"
	/codon_start=1
	/db_xref="Gene:600-119-124"
	/protein_id="AA52484.1"
	/db_xref="GI:182803"
	/translation="MQLELSTCFPLCLIRFCFSATRRYLGAVELSDMYQSDGLGELPVDAPRPVPKSPFENPVVYKTKLFEVYTHLPIINAKPRPPWMLGLPTQAEVYDTVVIILKMAHSPVLIHAGVSYKMADEAEVDDTSOREKDEKDFPGSSHTYVQVYLKENPMSADPCLITLYSYLSHVDLYKLSNGLIGALLCREGSLAKKTKTQHLKLIILFAVPEKGSMSHSEKTSNLSIMORDASASAMPKMHVNCVYNSRLGLIGCRKSYHHVIGKMTPEVSEITFLEGHTFLVDRNHRQASLEISPTTFLPAOTLMDLQGLFCHISHOHDEAEVAVKVDSCPEEPQLMKNNREAEVDDDTJSDMEVYAEFDDNSPSFLIARYVAKHNPRTWYHIAAEEDMDYAPVLPADRSYQYLNNQDRIKTKKAFKAYTDEFPTKTAIOHSGSLIGLPEVGDVTLIRKNOASRYNPIPIGIDVPLRYSLRPLGVKHLKDEPILPGEIFEIKWTVAIVDEDPSTSDRCILTRYSSVNNPRLQASDLIGPLLCKKESVDORGNQIMSKDRNVLLFSPDNRESVYLTENIOTPLPACQVLEDPFASNMHSINGVYDSLDLSCLAEVATWYILSLGAOTDFLSVFSCTYFKKAVYLEDVLLPEFSGETGFWMSMSPNGMLILGCHNSDRNKGKATALKVSSCDNGDYEDSYEDISAVLISKNNAIIEPRSPNSRHPSTROKOPNATIPENDIEKTDWFAHRTMPKIDNVSDDLMLRQSPHPGLSLDLQAKETPSDDPSGADISNNLSMTNHRPDLHSGDMVETPESGLOLRLENKLGITTAETELKLDKPYVSTNNLSTPSDNLAACTNISTGPPSPMPVHYDQLOTLTGKSSPLTESGPLSEENDSKLESGMLNSQSSMGKNVSTSESGHRLGKRAHGEPALITKONALFKVISILKTKNTSNNATRKTHIDGSLIENSPWQNLLESDETFKATVLIHOMLMDKNATLRHNSNKTTSKSMENYQOKEGRIPEPDAGNPMSEFKMLTLPESARWIDRTGKNSLGGSPSKQLVLEPERKSVEGONFLSEKKNVVGGEFTDGLKENYFSNRNLTNNLNLHENNTNHOEKKIOEIEKKEKTLIDENVLPQIHTVGTGKNMKNLFLSTRWDESYBGA YAPLOPRLSINDSTNTRKTHAFPSKSGEENLEGLNQTQKIVERYACTTRISPT SQQNFVORSKRALKQRLPLEETELERKTIIVDTSTQWSNMKHLRPLTQIDVNE KEKATIQSPSLDCLTNRSHSIPQANRPLPAKVSFSPSIRPIYLRVLRQDSNHPV AASARKRDSGVQESHSFLQGAKKNNLSLAILTEMTDQREVGSLGANSYTKVY ENYVLPKDEPKTSGKVELPKYHIYOKDLEPTSGNSGQVLDVGLSGLODEGAIKWNANRPGKVPRLRVATNESAARPSKPLRLPMDNMGYQIIPKEMKSOEKPERTA FKKDITLISLACSNNAIALINGQKKPELEYMAQOGRTELCSNPVYLKHORE ITRTTQLSDOEIYDDTISVMKKEFDYIDEDENSPSFOKTRHNYIAAVERIAM DYSGSSPHVLRNRAQGSVQPKKVFQVFOETDTSFQPIYRGLENLMDLPLPIA EVDENIWTGYNQASRPYSFSLISYEDROCAEARKNVFNENKTYFMKYOHIM APTEDEFDCKMAVEVDYDEKDVHSGSLIGLIVCHNNTLNPAHGROVYQDALFT IEDTKSWYETENNERNCRAPCQNTQMDPPPKENYRHALINGTINMOTLPYVAOOR IRTWLSMGNENHSTHFSGHVETVRKKEEYKALVNLPGVEFEVEMLPKAGTMR VECILGHLNAGMTLPLVSNKQTPPLGASGAIIRPOFTTASQVQMPKLARHY SGTNSAMSTKEPESMIKIVDLAPMIINGITQGRKQFSSLYISOFTIYSLDGKQO TYRNSNTGTLKMPFNGVDSGIIKINIPITIAIYLHPTIYISRTLEMLMGDL NSCGMPKMEKASIDAOITYASSFTMAMTAPWSKARLLOGRSNAWROVANNPKM LQNDQCTMKYGTGVTTCGVKSLTSMYKVEFLISSDQGHQMTFLPQNGKVKYFQNG DSEFPVNASLDPLRLTRILRLRHDSVWHDALRREYVLCERADLY"
BASE COUNT	2853 a
ORIGIN	160 bp 5' to Sacti cut site; chromosome Xq28.

[illegible]



QY 1021 agcttatgcagaagtagacaagctgtccagaggaaccccaactacgaatgaaaaataatga 1080  
|||||  
Db 1170 AGCTTATGTCAAAGTATAGACACTCTCCAGAGAACCCCACTACGAATGAAAAATTAATGA 1229  
QY 1081 agaagcggaaagactatgatgatgtactgtactgtcgtgaatgtggtgtgagtttga 1140  
|||||  
Db 1230 AGAAGCGGAAGACTATGATGATGATCTTACTGATTTGGAATGATGATGGTGCAGTTTGA 1289  
QY 1141 tgaatgaacactctccctcccttaccatctgcctcaattgtgcagaagaacacccaac 1200  
|||||  
Db 1290 TGAGTGAACACTCTCTCTTATCAAAATTCGCTCAGTTGCCAAGAAGCATCTCTAAAC 1349  
QY 1201 ttgggtacattacattgtctgtcgtgaagaaggagacttggaactgtcccttaagtcctcgc 1260  
|||||  
Db 1350 TTGGGTATACATTACATCTGCTGTAAGAGAGAGACTGGGACTATCTCTCTTAAGCTCGC 1409  
QY 1261 ccccgatgcacagaagttataaagtcacatttgaacaatgcccctcagcggatltgtag 1320  
|||||  
Db 1410 CCCCAGATGACAGAGATTATAAAGTCAATATTGAACAATGGCCCTCAGCGATTGGTAG 1469  
QY 1321 gaagtacaaaaaagtcggattttagtcatacacagatgaacactttaagactgtgaagc 1380  
|||||  
Db 1470 GAAGTACAAAAAAGTCGATTATGATACACAGATGAACCTTTAAGACTCGTGAAGC 1529  
QY 1381 tattcaagatgaatcagaatcttggaccttacttaltgagggaagltgagacacact 1440  
|||||  
Db 1530 TATTACACATGAATCAGAGATCTTGGGACTTTTATGGGAASTTGGAGACACACT 1589  
QY 1441 gtgatattatlttaagaatcaagaagcagaacacataaacatctaccctcagcgaatcac 1500  
|||||  
Db 1590 GTTGATTATATTAAAGATCAAGACAGACACATATACATCTACCCCTCAGCGAATCAC 1649  
QY 1501 tgaatgcgtcccttgtatcacaaggagatacccaaaagtgtlaaacaatttgaagaattc 1560  
|||||  
Db 1650 TGATGTCCTGCTTGTATTCAAAGAGATTACCAAAAGGTGAAMACATTTGAAGGATTT 1709  
QY 1561 tccattctcgcagagaagaataatccaataaaatgacagtgactgtagaagaatggcc 1620  
|||||  
Db 1710 TCCAAATCTGCGAGAGAAATATTCAATATTAATGACAGTGTAGTGGGCC 1769  
QY 1621 aactaaatcagatccctcgtgctcgaacccgtatctactagttcgttlaatatgagaag 1680  
|||||  
Db 1770 AACTAAATCAGATCTCGGAGCTGACCCGCTATTACTAGTTCTGTATATATGAGAG 1829  
QY 1681 agatctagcttcaggactcaattggccctctccatctgtctacaagaatctgtagatca 1740  
|||||  
Db 1830 AGATCTAGCTTCAGGACTCATTTGCCCTCTCTCATCTGCTACAAAGAACTGTAGATCA 1889  
QY 1741 aagaggaacacagaataatgtcagaacaagaagatgtcatcctgtttctgtatttga 1800  
|||||  
Db 1890 AAGAGGAACCAAGTATATGTCAAGCAAGAGAGATGTATCTCTTTCTGTATTTGATGA 1949  
QY 1801 gaaccgaagctgtacctcacaagaaatalacaagctttctcccaatccagcttgagt 1860  
|||||  
Db 1950 GAACCGAAGCTGTGACTCTCACAGAGATATACAACGCTTTCTCCCAATCCAGCTGAGT 2009  
QY 1861 gcagcttgaagatcccaagatccaagctccaacatcatgcaagcatcaatgtgctatgt 1920  
|||||  
Db 2010 GCAGCTTGAGGATCCAGATTCCAAAGCTCCAAATCATGACAGCATCAATGGCTATGT 2069  
QY 1921 ttctgataagcttgcagctgtcagcttcttgcataaggtgcatatcgtgtacattcgaag 1980  
|||||  
Db 2070 TTTTGTAGATTGTCAGTGTGTCAGTTTGTTCATGAGGTGSCATACTGTGATCTTAAG 2129  
QY 1981 cattgagcacagactgaactccttctgtctctctctcgtgatataccttcaaacacaa 2040  
|||||  
Db 2130 CATGTGAGCACAGACTGACTCTCTTCTGTCTCTCTGTGATATACCTTCAAAACACAA 2189  
QY 2041 aatgtgtctatgaagacacactcaacctatcccatctcagaagaacgtctctcatgtc 2100  
|||||  
Db 2190 AATGTCTATGAAGACACACTCAACCTATTCATCTCAGGAGAAACTGTCTTCATGTCTC 2249

QY 2101 gatgaaaccacgaagctcatgattctgggtgtgcacaactcgaacttccgaaacagag 2160  
|||||  
Db 2250 GATGGAACCCAGAGTCTATGATTTCTGGGGTGGCCAACTCAGACTTTGGGAACAGAG 2309  
QY 2161 catgaccgcttactgaagttctctagttgtgacaagaacactgtgtattatcagaga 2220  
|||||  
Db 2310 CATGACCGCTTACTGAAGTTCTAGTGTGACACAGAACCTGTGATTATATGAGAGA 2369  
QY 2221 caattatgaagatattcagacatactgtcgtgagttaaaaaaacaatgtcaatgaacaaag 2280  
|||||  
Db 2370 CAGTTATGAAGATATTTCAACATACCTGTGAGTAAAAACAATGCCATTTGAAACCAAG 2429  
QY 2281 ctctcccaaatccaagacacccctagcactgaagcaaaagcaatttaatgcaaccaca 2338  
|||||  
Db 2430 CTCTCCAGATTCAGACACCTTAGCCTAGGCAAAAGCAATTTATGCGCACACA 2487

Search completed: November 18, 2001, 01:27:15  
Job time: 29577 sec



**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2001, 13:22:11 ; Search time 27.33 Seconds

(without alignments)  
3263.004 Million cell updates/sec

Title: US-09-689-430-2

Perfect score: 7869  
Sequence: 1 MQELSTCFELCLRFCSA.....VWHLALRMVELGCEADLY 1471

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing:

Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
1: A.Geneseq\_0601.\*  
2: /SIDSR/gcgdata/geneseq/geneseq/AA1980.DAT.\*  
3: /SIDSR/gcgdata/geneseq/geneseq/AA1981.DAT.\*  
4: /SIDSR/gcgdata/geneseq/geneseq/AA1982.DAT.\*  
5: /SIDSR/gcgdata/geneseq/geneseq/AA1983.DAT.\*  
6: /SIDSR/gcgdata/geneseq/geneseq/AA1984.DAT.\*  
7: /SIDSR/gcgdata/geneseq/geneseq/AA1985.DAT.\*  
8: /SIDSR/gcgdata/geneseq/geneseq/AA1986.DAT.\*  
9: /SIDSR/gcgdata/geneseq/geneseq/AA1987.DAT.\*  
10: /SIDSR/gcgdata/geneseq/geneseq/AA1988.DAT.\*  
11: /SIDSR/gcgdata/geneseq/geneseq/AA1989.DAT.\*  
12: /SIDSR/gcgdata/geneseq/geneseq/AA1990.DAT.\*  
13: /SIDSR/gcgdata/geneseq/geneseq/AA1991.DAT.\*  
14: /SIDSR/gcgdata/geneseq/geneseq/AA1992.DAT.\*  
15: /SIDSR/gcgdata/geneseq/geneseq/AA1993.DAT.\*  
16: /SIDSR/gcgdata/geneseq/geneseq/AA1994.DAT.\*  
17: /SIDSR/gcgdata/geneseq/geneseq/AA1995.DAT.\*  
18: /SIDSR/gcgdata/geneseq/geneseq/AA1996.DAT.\*  
19: /SIDSR/gcgdata/geneseq/geneseq/AA1997.DAT.\*  
20: /SIDSR/gcgdata/geneseq/geneseq/AA1998.DAT.\*  
21: /SIDSR/gcgdata/geneseq/geneseq/AA1999.DAT.\*  
22: /SIDSR/gcgdata/geneseq/geneseq/AA2000.DAT.\*  
23: /SIDSR/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7857	99.8	1471	18	AAW23414
2	7775	98.8	1457	20	AAW21675
3	7771	98.8	1457	19	AAW46246
4	7771	98.8	1457	19	AAW44372
5	7674	97.5	1438	21	AAW21262
6	7642.5	97.1	1440	12	AAW12971
7	7616	96.8	1661	18	AAW18670
8	7615.5	96.8	1516	9	AAW80265
9	7585	96.4	1424	22	AAW48842
10	7574	96.3	1425	9	AAW80267
11	7571	96.2	1424	9	AAW80268

12	7567	96.2	1424	10	AAW91169	Sequence of 740 Ar
13	7416.5	94.2	2342	18	AAW11422	Active Factor VIII
14	7416	94.2	2351	18	AAW10591	Factor VIII:C (Phe
15	7416	94.2	2351	18	AAW13496	Factor VIII:C (Arg
16	7416	94.2	2351	20	AAW21676	Factor VIII protei
17	7415.5	94.2	2344	18	AAW11432	Active Factor VIII
18	7415.5	94.2	2344	18	AAW11410	Active Factor VIII
19	7415	94.2	2345	18	AAW11415	Active Factor VIII
20	7414.5	94.2	2346	18	AAW11421	Active Factor VIII
21	7414.5	94.2	2346	18	AAW11431	Active Factor VIII
22	7414.5	94.2	2346	18	AAW11434	Active Factor VIII
23	7414	94.2	2347	18	AAW11411	Active Factor VIII
24	7414	94.2	2351	18	AAW10592	Factor VIII:C (Tyr
25	7413.5	94.2	2348	18	AAW11409	Active Factor VIII
26	7413.5	94.2	2348	18	AAW11414	Active Factor VIII
27	7413	94.2	2349	18	AAW11424	Active Factor VIII
28	7413	94.2	2349	18	AAW11428	Active Factor VIII
29	7413	94.2	2349	18	AAW11430	Active Factor VIII
30	7413	94.2	2349	18	AAW11420	Active Factor VIII
31	7412.5	94.2	2350	18	AAW11413	Active Factor VIII
32	7412	94.2	2351	6	AAW50059	Human factor VIII.
33	7412	94.2	2351	18	AAW10590	Factor VIII:C (Phe
34	7412	94.2	2351	18	AAW11425	Active Factor VIII
35	7412	94.2	2351	18	AAW11427	Active Factor VIII
36	7412	94.2	2351	18	AAW11435	Active Factor VIII
37	7412	94.2	2351	18	AAW11408	Active Factor VIII
38	7412	94.2	2351	18	AAW11416	Active Factor VIII
39	7412	94.2	2351	18	AAW11419	Active Factor VIII
40	7412	94.2	2351	18	AAW11329	Native Factor VIII
41	7412	94.2	2351	19	AAW46245	Human factor III P
42	7412	94.2	2351	19	AAW44373	Human Factor VIII.
43	7412	94.2	2351	21	AAW52537	Human full-length
44	7411.5	94.2	2352	18	AAW11423	Active Factor VIII
45	7411.5	94.2	2352	18	AAW11426	Active Factor VIII

#### ALIGNMENTS

RESULT 1	
AAW23414	standard; Protein: 1471 AA.
AC	AAW23414;
DT	08-APR-1998 (first entry)
XX	
DE	Human B-domain deleted factor VIII protein.
XX	
KW	Post-translational regulatory element; PRE: enhancer II; intronless gene;
KW	surface antigen gene; cytoplasmic accumulation; targeted delivery;
KW	near consensus splice sequence; blood coagulation factor; Factor VIII;
KW	factor IX.
XX	
OS	Homo sapiens.
XX	
PN	W09733994-A1.
XX	
PD	18-SEP-1997.
XX	
PF	10-MAR-1997; 97MO-U0503561.
XX	
PR	11-MAR-1996; 96US-0683839.
XX	
PA	(IMMU-) IMMUNE RESPONSE CORP.
XX	
PI	Bidlingmaier S, Ill CR;
XX	
DR	WPI: 1997-470874/43.
XX	
DR	N-FSDB; AAT73164.
XX	
PT	Vector for increased expression of intronless genes - comprises
XX	intronless gene with at least one near consensus splice sequence, a



PT promoter and at least one viral cis-acting post-transcriptional  
PT regulatory element  
XX  
PS Example 1: Pages 31-36; 59pp; English.  
XX  
CC The present sequence represents human B-domain deleted factor VIII  
CC the cDNA encoding this protein also contains, 3' of the coding region, a  
CC post-translational regulatory element (PRE) of the Hepatitis B virus.  
CC PRE sequences have been shown to function in cis to increase the  
CC steady-state levels of surface gene transcripts by facilitating  
CC cytoplasmic accumulation of these transcripts. The above nucleic acid  
CC sequence is part of a novel vector, comprising an intronless gene  
CC containing 1 or more near consensus splice sequences operably linked to  
CC a promoter sequence so that the gene is transcribed in a cell.  
CC Intronless gene transcripts which contain near consensus splice site  
CC sequences are believed to get tied up in the nucleus of the cell where  
CC splicing occurs, rather than being transported to the cytoplasm where  
CC they can be translated into proteins. The PRE sequences are transcribed  
CC along with the gene, causing export of the gene transcript from the  
CC nucleus into the cytoplasm of the cell. The vector can be used  
CC to increase the expression of an intronless gene containing at least one  
CC near consensus splice sites, preferably cDNA encoding a blood coagulation  
CC factor, particularly factor VIII or IX. The complex allows the targeted  
CC delivery of the vector to a specific cell, e.g. hepatocytes when the  
CC ligand is an asialoglycoprotein which binds the asialoglycoprotein  
CC receptor present on their surface.  
CC  
CC  
CC  
XX  
XX  
SQ Sequence 1471 AA:  
  
Query Match 99.8%; Score 7857; DB 18; Length 1471;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1469; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
QY 1 MQELSTGCFECLRFCSATRRRYLGAVELSMDYMSDLGLPVDARPPRPKSPFN 60  
DB 1 mvelstcfcllrfcsatrrrylgaVELSMDYMSDLGLPVDARPPRPKSPFN 60  
DB 1 tsVVYKKTLEFEFTVHLEFNIAKRRPPMGLGPTTQAEVYDVTITLKMASHVSLNAV 120  
QY 61 TSVVYKKTLEFEFTVHLEFNIAKRRPPMGLGPTTQAEVYDVTITLKMASHVSLNAV 120  
DB 61 tsVVYKKTLEFEFTVHLEFNIAKRRPPMGLGPTTQAEVYDVTITLKMASHVSLNAV 120  
QY 121 GVSITWKSSEGAEYDQTSQREKEDKVPFGSGHTYVQYLKNGPMASPLCLTITSYLSH 180  
DB 121 gvsitwkssegAEYDQTSQREKEDKVPFGSGHTYVQYLKNGPMASPLCLTITSYLSH 180  
QY 121 gvsywkasegeayddqtqrekdckvfpqsgshcyvqylkenqmaspRLcltySLsh 180  
DB 121 gvsywkasegeayddqtqrekdckvfpqsgshcyvqylkenqmaspRLcltySLsh 180  
QY 181 VDLKDNLSGLIGALLVREGSLAKEKQTLLKFTLLFAVPEEGSMHSETNSLMODRD 240  
DB 181 vdlkdnlsgllgallvcregsIakeKcqlhklfllfavldeqkswnsetknsImqdrd 240  
QY 241 AASARAMPKMTVNGVYVNSLPGLLGCHRRKSVYWHVIGMTTPEVHSIFLEGHTFLVRNH 300  
DB 241 aasaraMpkmtVngvYvNslpGllgchrrksVYwhvIgmTtpEvhsIfleghtflvrnh 300  
QY 301 RQASLEISPTFLRAQTILMDLGFLLFCHTSSHOHOGMEATVKKVDSCEPEEQLLMKKNNE 360  
DB 301 rqaSleIsptflraQtiLmdLgFllfChTsshOhOGmeaTvKkVdscePeEqllmKknNe 360  
QY 361 EAEYDDDLTSEMDVYVFEFDDNSPFIQIRSVAKKHPTWVHYTAAEEDMDVAPLYLA 420  
DB 361 eaeYdddlTsemdvYvFefddnsPfiQirsvAKKhptWvHyTaaEedmdvApLyLa 420  
QY 421 PDDRYSKQYLNNGPQIRGRRYKRVFMAVYDTEFPKTRBATQIHESGILGPLLYGEGDTL 480  
DB 421 pddrYSkqYlNngpqIRgrrYkRVfMaVYdTEfPKtRbaTqIHesGIlgPlLYgeGdtL 480  
QY 481 LIIFKNOASRPYNYIPHGITVVRPLYSRRLKGVNHLKDFPLLPGLIIRKYKATVVEDGP 540  
DB 481 lIIfkNoasrPYnYIPhgITvvrPLYsrRLkGVnHLkDfPLlPGLIIRkyKaTVveDGP 540  
QY 541 TKSDPRCTRYSSPVNNERDLASGLIGPLITCYKESVDORGNQMSOKRANVILTSVPE 600  
DB 541 tkSDprCTrYsspVnNeRdlASglIGPlITcyKESvdORGNqMSOKraNVilTSvPe 600

QY 601 NRSWYLTENIQRELPNPAVOLDEDEFOASNMHSINGVFDLSQLSVCLHEVAYWYLS 660  
DB 601 nrsWylTeNiqrElPnpagVqldeDeFoaSnmHsIngvFdlsQlsvclheVayWyls 660  
QY 661 ICAQDFLSVPSFGYTFKHKMYEDTLTLFPSSGTVMSHENGIMTLGCHNSDFNRNG 720  
DB 661 iCaQdfLsvpsfgYtfKhkMYedTLtlfPssgTvmsHengImtlgChnsdfnrng 720  
QY 721 MWALKVSSCDNTDDEYEDSYEDISAVLLSKNNAIPEERSONSRRHPSTOKOFNATPP 780  
DB 721 mWAlkvsscdntdDeYedSyedIsavllSknnaiPeersonsRRhpstOkofNatpp 780  
QY 781 VLKRHOREITRTTLOSDEIDYDPTISVENKKEDEFDIYDEENOSPRSFQKTRHYFA 840  
DB 781 vlKrhoreITrttLosdeIdyDptIsvenKkedFDiYdeEnosprsfqKtrhyfia 840  
QY 841 AVERLMDGMSSSPVNLNRNRAOSGSVPQKVVYRQEFIDGSEFTQPLVNGELNEHLGLLP 900  
DB 841 aveRlmdgmssspVnlNrnrnRAosGSvPqKvvYrQefIdgseftQplvngelnehlglLP 900  
QY 901 YTRAEEEDINWTFRNQASRPYSFSSLSIYEEDROGAEPKXNFVKNPRTKTFEMKVOH 960  
DB 901 ytraeeedInwTfrnQasrPyssfssLSIyeedRoGaepKxnfVknPrtkTfemKvoh 960  
QY 961 HMAPTKDEFDCKAMAYFSDVDLEKDVHSGLIGPLLVCHTNTLNPAGRQVTVQEFALFT 1020  
DB 961 hmaPtKdeFdckAmayfSDvdLEkdvHsglIGPlLVchTntLnpagRqvTvQefalft 1020  
QY 1021 IFDERKSWYFTFNMERNRACPNQIOMEDPTFRKENTRFAINCYIMDTPLGVLMAQDORIR 1080  
DB 1021 iFderKswyftFnmErnrAcPnqiOMedptFRkEntRfaIncyImdtPlgvlMaqDorir 1080  
QY 1081 WYLLSMGSENENHSIHFSGHVTVARKKEEYKMALYNLPGVEFEVEMLPKAGIMRECL 1140  
DB 1081 wyllsmgsenEnhsIHfsgHvtvArkkeeYkMaLynlPgveFeVemlpKagImrecl 1140  
QY 1141 IGEHLHAGNSTLELVYSNKCQTPILGMASGHTRDQITASGOYGWAPKRLARLHYSSTINA 1200  
DB 1141 igehlhagnstlElvysnKcqpIlgmASghtrDqitAsgoYgWaprkLarLhysstina 1200  
QY 1201 WSTKEPFSIKYDLAPMHINIGIKTGAROKFESSIXISQFIIMYSLOCKKQOTYRGSTG 1260  
DB 1201 wstkePfsIKydlAPmhInigIKtgARokfESSIXISqfiImysLoCKkqotYrgstg 1260  
QY 1261 TLMVFEQVNDSSGICKHNFNFPITIRYTRLPTNHSISTRLMELMGDLMSCMPRLGME 1320  
DB 1261 tlmvfeqVndssgIkhnfnFPitIRytrLpTnhsistRLmElmgDLmscMprlgme 1320  
QY 1321 SKAISDAQITTASSYFTNMPATWSPSKARLHLQGRSNAMPQVNNPKEMLOVDFQTKMYKT 1380  
DB 1321 skaisdaQitTassYftnMpatWspSkarLhlQgrsnAmPqvNnpkEmloVdfqtkmYkt 1380  
QY 1381 GVTGQVKSLLTSMVYKFEELISSODGHOMTLFPONGKVVYKFOGNODSFPTVYVSLDPL 1440  
DB 1381 gvtGqVksllTsmvYkfeELissodGHomtlFPongkVvYkfoGnoDSfptVvslDpl 1440  
QY 1441 LFRYLRHPQSVWQHIALRMEVLGCEADOLY 1471  
DB 1441 lfrYlrhpqsvWqhIalrMeVLgceAdoly 1471  
  
RESULT 2  
AAY21675  
ID AAY21675 standard; Protein: 1457 AA.  
XX  
XX AAY21675;  
AC  
XX  
XX  
DT 18-AUG-1999 (first entry)  
XX  
XX Beta-domain deleted Factor VIII protein.  
DE  
XX  
KW Factor VIII protein; gene modification; gene therapy; clinical disorder;



KW splicing pattern: RNA processing; gene regulation; beta-domain; human.  
XX Homo sapiens.  
XX MO9929848-A1.  
XX 17-JUN-1999.  
XX 25-NOV-1998; 98WO-US25354.  
XX 16-JAN-1998; 98US-0071596.  
XX 05-DEC-1997; 97US-0067614.  
XX (IMMU-) IMMUNE RESPONSE CORP.  
XX  
XX Biddingmaier S, Gonzales JEN, Ill CR, Yang CQ;  
XX WPI; 1999-385602/32.  
XX N-PSDB; AAX82258, AAX82259, AAX82260.  
XX  
XX Genes and vectors exhibiting increased expression and novel splicing  
XX patterns, useful for expression of, e.g. beta-domain deleted factor  
XX VIII  
XX  
XX Disclosure: Page 72-78; 123pp; English.  
XX  
XX The invention describes novel genes and vectors exhibiting increased  
XX expression and novel splicing patterns. It provides a gene encoding a  
XX Factor VIII protein, that comprises one or more consensus or near  
XX consensus splice sites which have been corrected to increase expression.  
XX The method, DNA sequences and expression vectors can be used to increase  
XX the expression of a gene, especially a Factor VIII gene. Genes containing  
XX modified 5' and/or 3' untranslated regions have optimized expression  
XX levels and tissue-specific expression. The methods are used for  
XX identification and correction of consensus splice sites, addition of  
XX introns, optimization of 5' and 3' untranslated regions and increase in  
XX cytoplasmic RNA accumulation. Hence the DNAs are useful in gene therapy  
XX to treat a clinical disorder, to study RNA processing and/or gene  
XX regulation. The present sequence represents a beta-domain deleted Factor  
XX VIII protein.  
XX  
XX Sequence 1457 AA:  
SQ  
Query Match 98.8%; Score 7775; DB 20; Length 1457;  
Best Local Similarity 99.0%; Pred. No. 0;  
Matches 1456; Conservative 1; Mismatches 0; Indels 14; Gaps 1;  
QY 1 MQEISNCFEFLCLDFCFSATRRYYLGAVELSDMYMOSDLGELPYDARFPRVPKSPFN 60  
DB 1 meletstcflcllrfctsatrryllgavelswdymgsdglpvdafiprvpkstfph 60  
QY 61 TSVVYKKTLEVEFTVHLFNIAKPRPPWMLGPTIQAEVYDTVYITLKNMASHVSLHAY 120  
DB 61 tsvvyykktlfeftvhlfniaakprppwmlgptiqaevydtvyltlknmashvslhav 120  
QY 121 GVSVMKASEGAEVYDDQTSQREKEDKVPFGSGSHYVQVLEKNGPMASDPLCTIYSYLSH 180  
DB 121 gvsvmkasegaevyddqtsqrekedkvfpqgshyvwqvlkengpmasdpclctysylsh 180  
QY 181 VDLVKDLSNGILGALLVCREGSLAKKETQTLHKEFTLLFAVDEKSMHSEKNSLMORD 240  
DB 181 vdlvkdlngsilgallvcregslakektqtlhkftllfavdegkswhseknslmqdird 240  
QY 241 AASRAMPKMHTVNGYVNRSLPGLIGCHRSVYVHVIIGMTPEVHSTIELGHTFLVNRH 300  
DB 241 aasrampkmhtvngyvnrslpgligchrsvyvwhviigmtpevhstieleghtflvnrh 300  
QY 301 ROASLETSPITFLTAQTLMLDGLFLFCHTSSHOHOGMEKVVAVDSCPEEPOLRMKNE 360  
DB 301 rgasletspitfltaqtlmldglflfchthssbhdgmeavkvvdscpeepqlrmkne 360  
QY 361 EAEDYDDLTDSEMDVVFDDDNSPSFTQIRSVAKKHPKTVWVYIAAEEDMDVAPLVLA 420

DB 361 eaedyddldtsemdvvrfdnnspstfiqirsvakkhpkvtwvhyiaaeedmdvapyvla 420  
QY 421 PDDRSYKSOYLNNNGFORIGRKYKKVRFMAVYDTEFKTBEALQIHESGILGPLLXGEVDTL 480  
DB 421 pddrsykesqylngngpqirigrkykkvrfmayudetfktrealqhesgilgplllxgevgdtl 480  
QY 481 LIIFRNOASREYNIRPHGITTVPRLYSRRLKGVNHLKDFPLPEGLTKYKWTYVVEGP 540  
DB 481 lliifrngasrpylnlyphgittvprlysrllkgyvnhldfpllpgeitkkykwtvvegpr 540  
QY 541 TKSDPRCLTRYSSPFWNNERDLASGLIPLLICYESVDGONGNOMSKRNVLTFSVDE 600  
DB 541 tksdprcltryssfwmnerdlaaglliplllcykesvdqngnmskdnvltfsvde 600  
QY 601 NRSWYLTENIGRFLPNPAGVOLDEDFQASINMHSINGVFDSTQLSVCLHEVAYWYLLS 660  
DB 601 nrswyltenigrflpnpagvoldebefqasnlimhsingvfdslqsvclhevaywylls 660  
QY 661 IGAQTDLFVSFSGYTFKKHKKVYEDTLTLFPFSGSTVPMSENPRGLMILGCHNSDFRRNG 720  
DB 661 lgaqtdlfvsfsgytfkhhkwyedtltlfpfsgstvmsemprglmlgchnsdfrrng 720  
QY 721 MTALLKVSQCDKNTGDYEDSYEDISAVYLLSKNNAIEPRFSQNSRHSBTRKOFNATPP 780  
DB 721 mtallkvsqcdknsgdyedsyedisayllsknnaieprstsqn-----pp 780  
QY 781 VLKRHOREITRTTQLSDQEEIDYDDTISVEKKKEDEDIYDEDENSPFSQKTRRHYFLA 840  
DB 781 vlkrhoreitrttqlsdqeeddyddttsvenkkkedfdydedenqspfsqktrhyfla 840  
QY 841 AVERLMQMGSSPPHVLNBRNOSGVPQFKRVYVQEFNDGSEFTQPLVXGELNENHGLLGP 900  
DB 841 averlmqmgsspphvlbrnosgvpqfkrvyvqefndgseftqplvxlghlnenhgllgp 900  
QY 901 YIRAEVEENINWTFRNOASRPYSFSSLSIYEEDROGAEPRKKNFVNKETKTFYFWKVOH 960  
DB 901 yiraeveeninwtfnrnoasrpysslsisyeedrogaeprrkknfvnketktyfwvkqh 960  
QY 961 HMAPTKDEFDCKAMAYFSDVDLEKDVHSGILGPLLCHTNTLNPAGHQTVQEFALFT 1020  
DB 961 hmaptkdefdckamayfsdvdlekdvhsgilgplllchntclnpahgqvteqefalft 1020  
QY 1021 IFDEFTKSWTFENENRNCRAQNTQMEDPTFKENVRFLAINGYIMDTPLGVLMMADQDQIR 1080  
DB 1021 ifdefkswtfenennrcraqntqmedptfkenvrflaingyimdtpkgvlmmadqdqir 1080  
QY 1081 WYLLSMGSNENIHSHFSGHVFTRKKKEEYKMALYNLPGVEFVEMLPKAGIWRVETCL 1140  
DB 1081 wyllsmgsnenihshfsgvhvtrkkkeeykmalynlpgvefvevmlpskaglwrvetcl 1140  
QY 1141 IGEHLHAGMSTFLVYNSKCCQTPLGMAAGIRDFQITASGOYGWAPKLARLHVSCTINA 1200  
DB 1141 igenlhagmstflvynskcqtplgmaaghrdfqitagsyggwqapklarlhvsctina 1200  
QY 1201 MSTKEPFWIKIVDLAPMIIHGITGQAROKFESSLXIQTIFIMSLDOCKKQYRGNSTG 1260  
DB 1201 mstkepfwmikivdlapmiihgittgqarokfesslsxiqtfimsldockkqyrgnsgtg 1260  
QY 1261 TLMWFFGNVDSGKIHNFNPFIARVRLAPHTVYSIRSTLRMELMGDGLMSCMPLGME 1320  
DB 1261 tlmwffgnvdsdgkihnfnfpfiarvrlaphtvysirstlrmelmgdglmscmplgme 1320  
QY 1321 SKAISDAOITASSYFTNNFATWSPSKARLHLOGSNAMARPOVNNPKEMLOVDFOKTMKVT 1380  
DB 1321 skaisdaoitassyftnnfatwspskarlhlogsnamarpoynnpekemlovdfoktmkvt 1380  
QY 1381 GVTTOGVKSLTSMVYKVEFLISSODGHQWTLFPONGKVKVFKQGNQDSFTPVVNSLDPL 1440  
DB 1381 gvttogvksltsmvykveflissodghqwtlfpongkvkvfkqgnqdsftpvvnslldpl 1440  
QY 1441 LTRYLRIHPOSWVHQAIRMEVLGCEAODLY 1471  
DB 1441 ltrylrhihposwvhqairmevlgceaodly 1471



Db 1427 ltrylrlnpgswvlnqialrmvlgceagdlly 1457

RESULT 3  
AAW46246  
ID AAW46246 standard; Protein: 1457 AA.  
XX  
AC AAW46246;  
XX  
DT 06-AUG-1998 (first entry)  
XX  
DE Human factor VIII beta-domain deleted SQN deletion protein sequence.  
XX  
KW Replication defective; recombinant retrovirus; RVV; therapeutic protein;  
KW haemophilia; thrombosis; hypercoagulable disorder; liver disease; human;  
KW hepatitis; thalassemia; phenylketonuria; Lesch-Nyhan syndrome; diabetes;  
KW cystic fibrosis; Duchenne's Muscular Dystrophy; hypercholesterolemia;  
KW hypopituitarism; adenine deaminase deficiency; HIV infection; anaemia;  
KW Gaucher's syndrome; high blood pressure; Alzheimer's disease; autoimmune;  
KW inflammatory disease; factor VIII.  
XX  
OS Homo sapiens.  
XX  
PN WC9800541-A2.  
XX  
PD 08-JAN-1998.  
XX  
PF 02-JUL-1997; 97WO-US11784.  
XX  
PR 04-JUN-1997; 97US-0869309.  
PR 03-JUL-1996; 96US-0645601.  
PR 13-AUG-1996; 96US-0696381.  
XX  
PA (CHIR ) CHIRON CORP.  
XX  
PI Allen JR, Barber JR, Boder M, Chang SMW, Chong K;  
PI De LA VEGA D, Depoloni, Greengard J, Hsu DC, Idanez CE;  
PI Jolly DJ, Lee R, Mittelstaedt DM, Prussak CE, Respass JG;  
DR WP1; 1998-086966/08.  
DR N-PSDB; AAV19581.  
XX  
PT New replication defective recombinant retro-viruses - which can be  
PT administered to provide long term systemic expression of therapeutic  
PT protein in blood, useful in, e.g. treating hyper-coagulable  
PT disorders  
XX  
PS Example 28; Pages 213-217; 272pp; English.  
XX  
CC This is the beta-domain deleted SQN deletion protein of human factor  
CC VIII. The encoding DNA is used to construct recombinant retroviral  
CC vectors expressing human factor VIII. The invention provides the  
CC preparation of replication defective recombinant retrovirus (RVV)  
CC expressing a therapeutic protein. The RVV preparation is resistant to  
CC degradation by human complement and is capable of inducing long term  
CC systemic expression of the therapeutic protein when administered  
CC intravenously to a human. The long term systemic expression results in a  
CC measurable level of the therapeutic protein being produced in the blood  
CC of the human for a period of at least 30 days after the administration of  
CC the RVV vector preparation. RVV's can be used for in vivo delivery of  
CC therapeutic protein to treat, e.g. haemophilia A, haemophilia B,  
CC thrombosis, hypercoagulable disorders, liver diseases such as hepatitis,  
CC disorders such as thalassemia, phenylketonuria, Lesch-Nyhan syndrome,  
CC severe combined immunodeficiency (SCID), cystic fibrosis, Duchenne's  
CC Muscular Dystrophy, inherited emphysema, familial hypercholesterolemia,  
CC diabetes, hypopituitarism, adenine deaminase deficiency, alpha-  
CC antitrypsin deficiency, Gaucher's syndrome, anaemia, infections such as  
CC HIV infection, high blood pressure, Alzheimer's disease, autoimmune or  
CC inflammatory disease or graft versus host disease. RVV's are capable of  
CC surviving inactivation in human serum thereby allowing efficient gene  
CC transfer over prolonged periods of time.  
XX  
Sequence 1457 AA:

Query Match		98.8%; Score 7771; DB 19; Length 1457;
Best Local Similarity		99.0%; Pred. No. 0;
Matches 1456; Conservative		0; Mismatches 1; Indels 14; Gaps 1;
QY	1	MOLEISTCFPLCLRFCSATRRYILGAVELSWDMQSDLGELPVDANFPPRVKSPFPN 60
DB	1	mqlelstcfllclrfcfsatrryylgavelswdmqsdlgelpvdarfprrvksfpfn 60
QY	61	TSVYVKKTLFEVEFTVHLFNIAPKRPMMGLGPTLOAEYDVDTVTLLNMASHPSLAV 120
DB	61	tsvyvkktlfevftdhlfniakprppmgliprtlgaeydvtvtllnmashpslhav 120
QY	121	GVSYWKASGAEYDDQTSOREKEDDKVPFGSGHTVWOLKENGPMASDPLCLTYLSH 180
DB	121	gvsywkasegayeddtsgrekdckvfpgshlyvwqlkengmaasdpclctysylsh 180
QY	181	VDLVKDLNSGLIGALLVCBESLAKETQTLAKFTLLRAVPDEGKSMHSETRKNSLMQDRD 240
DB	181	vdlvkdlnsgllgallvcrgeslakektqlhkflllfavidegkswshetknslmqdrd 240
QY	241	AASARAMPKMHITVNGVYVNSLPLGLGCHRSYVWHVIGMTPEVHSIFLEGHFLVFNH 300
DB	241	aasaraampkmhltvngvynslpjlighchksywhvigmtphevhsifleghtflvfnh 300
QY	301	ROASLEISPTFLTAQTLMDLGOFLFCCHISSHOHDGMEAVYKVSCEPEPOLMKNE 360
DB	301	rgasleisptfltaqtlmdlmgflfchishqhdgmeaykvkdscepepqrlmkne 360
QY	361	EAEEDDDDLTDEMDVYRRDDNNSPFIQIRSVAKKHRTWVHTAAEEDMDTAPLYLA 420
DB	361	eaedydddltdsemdvrrdddnspfiqirsvakkhpkrtwhtyaaeeedwdyaplyla 420
QY	422	PDDRSTKSOYLNNQPORIGRKRYKKVFPMAYPDETFKTRTALIOHESGILGPLYGVGDTL 480
DB	421	pdrrstksylnnqporlgrkrykkvrfmaypdectfktreaiohneqilgpllyvgvdtl 480
QY	481	LILFFKQASRPYNIAPHGITDVRPLYSRRLPKGVKHLKDFPLIPGEIRKMYTAVDEGP 540
DB	481	lilffkngasrplyniphgltvdrplysrllpkgvkhlkdfplipgeikmytavegpr 540
QY	541	TKSDRCLTRYSSYVNNERDLASGLIGPLICYESVDQGNQMSKRVNLLFSVDE 600
DB	541	tksdprcltryssfvnnerrdlaasgligpllcykesvdqgnqmsdkrvnllfsvde 600
QY	601	NRSWITENIOFELPNPAGVOLEDEPEFOASNTMHSINGVFPDLSQVCLHEVAYWYLLS 660
DB	601	nrswltenioqflpnpgvqledepefasnltmhsingvfdlsqsvclheavwylls 660
QY	661	IGAQTDFLSVFFSGYTFKKHWYEDTLTLPFSGEFTVMSMENPGLWILGCHNSDFRNRG 720
DB	661	igqtdflsvffsytffkhwmyedtlclpfsgctvmsmenpgllwlgchnsdfrnrg 720
QY	721	MWALLKVSCKDNKGTYEDYEDISAYILSKNNALIEPRFSQNSRHSSTQKOFNATPP 780
DB	721	mwallkvsscdnktgyedsyedisayilsknnaleprfsqn-----pp 780
QY	781	VKRRHORETTRTTLSDOEIDYDPTISVEMKKEPDFIYDEENSPSPFOKTRHYPIA 840
DB	767	vlkrrhorettrtltsdoeidydpdtisvemkkepdfdyedenpspsfkkcthyfia 826
QY	841	AVERLMDYGMSSSPVLRNRAQSGSVPOFKKVPFOEFTDGSFTOPLYRGELNENHGLGLP 900
DB	827	averlwdygmssspvlrnraqsgsvpqfkvrvfgetfdsgftplpyrgelnehlgllypr 886
QY	901	YRAEVEDINIMTFRNQASRPYSFYSSLSIYEEDRGCAEPRKNFVKNETKTYTFKWKQH 960
DB	887	yraevedimimtfngasrpyfsfyslsyeedrggaepkrfvkncktyfkvwqh 946
QY	961	HNAPTKDEPDCKAMAYFSDVLEKDVHSGILGPLLYCTNTLNAHGNQVVOFEALFET 1020
DB	947	hnaptkdefdckawayfsdvlekdvhsgllgpllvctntclnphagrqvltvgefalft 1006



QY 1021 IFDETRKSWFTEENMERNCRAPCNIOMEDPTEKENYRFAHANGYIMDTPLGLVMAODRIR 1080  
| | | | |  
Db 1007 lfdetksyrfteemerncrapenimedptfkenyrfinaingyindt.pglvmaodqrir 1066  
QY 1081 WYLSMGSNENIHSHSGHVFTRKKEEKYKALNLYPGVEYEMLPKAGIWRVCEL 1140  
| | | | |  
Db 1067 wylismgsnenishsfghvftvrkkeeekmalynlypyfetyemlpkagivrvec1 1126  
QY 1141 IGEHHAAMSTFLVYSKCCQPIGMASGHTRDQITASGOYGQAPKRLARLHYSGSINA 1200  
| | | | |  
Db 1127 lgehhaamstflvyskccqpiymasghtrdqtasgyqgwapkrlarlhsysina 1186  
QY 1201 WSTKEPFMIKVDLAPAIHIGIKQAGAROKFSSLYISQFIIMYSLDCKKQTYGNSTG 1260  
| | | | |  
Db 1187 wstkepfmikvdlapahigikqagarkfsslyisqfiimysldckkqtygnstg 1246  
QY 1261 TLMVFGVNDSSGIRKHNIFNPPIARVIRLHPHYSIRSTLRMELMGCDLNSCMIPLGME 1320  
| | | | |  
Db 1247 tlmvfgvndssgirkhnifnppliarvirrlhphysirstlrmeimgcdlnscmplgme 1306  
QY 1321 SKAISDAITASSYFTNMFATWSPSKARLHOGRSNMRPOVNNKEMLOVDFQTKMYT 1380  
| | | | |  
Db 1307 skaisdaITASyftnmfatwspskarlhgrsnmrpovnnkemlovdfoqtkmyt 1366  
QY 1381 GVTTOGVSLTSMYKREFLISSODGHQWTLFQNGKVKYFQGNODEFTPVNSLDPEL 1440  
| | | | |  
Db 1367 gvttogvsltsmykreflissodghqwtlfqngkvkyfognodeftpvnsldppl 1426  
QY 1441 LTRYLRHPQSWHQAIRMEVLGCEADOLY 1471  
| | | | |  
Db 1427 ltrylrhpgswbqialrmevlgceadqly 1457

## RESULT 4

AAW4372  
ID AAW4372 standard; Protein: 1457 AA.

AAW4372:

20-JUL-1998 (first entry)

Human Factor VIII SON deletion mutant.

Factor VIII: blood clotting; haemophilia A: gene therapy;

retrovirus; vector; human.

Homo sapiens.

Synthetic.

WO9800542-A2.

08-JAN-1998.

02-JUL-1997; 97WO-US11785.

04-JUN-1997; 97US-0869309.

03-JUL-1986; 96US-0645601.

13-AUG-1996; 96US-0696381.

(CHIR ) CHIRON CORP.

WPI; 1998-086967/08.

N-PSDB; AAV15338.

New replication defective recombinant retroviruses - which express B

PT domain-deleted human factor VIII or human factor IX for the

treatment of haemophilia

PS Claim 5; Page 175-180; 236pp; English.  
XX This polypeptide comprises the B domain deletion mutant SON of  
CC human Factor VIII. The SON mutant is created by fusing Ser-743 to  
CC Gln-1638 of native Factor VIII (see AAW4373) to form a Ser-Gln-Asn  
CC (SON) link between the A2 and A3 Factor VIII domains. A DNA  
CC sequence encoding the SON deletion mutant is provided in AAW15338.  
CC When compared to plasmid-derived Factor VIII, the SON deletion does  
CC not influence the in vivo pharmacokinetics, but the reduced size of  
CC the molecule appears to decrease proteolytic degradation. The  
CC invention relates to preparations of replication defective  
CC recombinant retrovirus (RV) expressing a B domain-deleted human  
CC Factor VIII protein, where the recombinant RV is capable of  
CC infecting human cells, is resistant to degradation by human  
CC complement and is capable of inducing long-term (at least 30 days  
CC and up to 6 months or longer post-injection) systemic expression of  
CC Factor VIII when administered to a haemophilia A patient.

SQ Sequence 1457 AA;

Query Match 98.8%; Score 7771; DB 19; Length 1457;  
Best Local Similarity 99.0%; Pred. No. 0;  
Matches 1456; Conservative 0; Mismatches 1; Indels 14; Gaps 1;

QY 1 MOIELSTQCFILRPFCSATRRYVLAVERLSMDVMOSDLGELPYDARFPPRPVSPFEN 60  
| | | | |  
Db 1 mqieltstqcfilrpfcsatrryvlavelswdvmgsdlgelpydarfprvpksfpfn 60  
QY 61 TSVVYKKTLEFEVTHLEFNIAKPRPPMGLGPTIOAEVYQTVVYTLKNMASHPSLNAV 120  
| | | | |  
Db 61 tsvyvkktlefevthlefniaakprppmglgptioaevyqtvtvylknmashpslnav 120  
QY 121 GVTYKASEGAEYDDQTSQREKEDKVPFGSHYVWOLKENGPMASDPLCLTYSTYSH 180  
| | | | |  
Db 121 gvytkasegaeyddqtsqrekedkvfpgshyvwvvlkengpmasdpclctystysh 180  
QY 181 VDLVKNLSGLIGALVYREGSLAKEKTOPLHKFTLFAVDEGKSMSEFNKSNLMOORD 240  
| | | | |  
Db 181 vdlvknlsiglalvycrgslakektpkhkftlfavdegksmseknsnsmord 240  
QY 241 AASARAWPMHTVNGVYVNSLPLGLGCHRSYVWHVIGMTTPEVHSIFLGGHFLVNNH 300  
| | | | |  
Db 241 aasarawpmhtvngvyvnsplplglgchrsyvwvhvigmtpvehsfllgghflvrnh 300  
QY 301 ROASLEISPIITFLRAQTLMDLGOFLFCHITSSHOHDMEAIVKVDSCPEEPOLRMKNE 360  
| | | | |  
Db 301 rgaaleispiitflraqtlmdlgqflfchissbqhdmeayvkvdsceepqlrmkne 360  
QY 361 EAEDYDDDLTDSMDVYVFEEDDNSPSFQIRSVAKKHKKTWVHYIAAEEDMDVAPLYLA 420  
| | | | |  
Db 361 eaedydddltdsmdvyvfeeddnspsfqiirsvakkhkwvhyiaaeedmdvayplyla 420  
QY 421 PDDRSYKSOYLNNNGFORIGRKYKKVRFMAVYDEFEKTRKATQIHESGILGPLYGEVGDTL 480  
| | | | |  
Db 421 pddrsykysoylngngforigrkykkvrfmayvdeftktrcatqihesgilgpllygevgdtl 480  
QY 481 LITRKNASRPYNTIPAGITDVRLYKRRLPKGYKHLKDPILPGEITFKYKWTYVEDGP 540  
| | | | |  
Db 481 litrknasrpyntipagitdvrlpkykrrlpkykhlkdpilpgeitfkykwtvvedgp 540  
QY 541 TKSDPRCLTRYSSFVNNERDLASGLIGPLLICKESVDGONGNIMSKRNVITFSVDE 600  
| | | | |  
Db 541 tkdprcltryssfvnnerdlasgligpllickesvdgongnimskrnvifsvde 600  
QY 601 NRSWLTENIQRFLEPNPAGVLEDEPEFQASNIMHSINGVYVDSQLSCLAEVAYWYILS 660  
| | | | |  
Db 601 nrswlteniqrflepnpagvledepefqasnimhsingvyvdsqslsclaevaywyils 660  
QY 661 IGAOTDLISVFFSGYTRKHKWYEDTLTLPPSGEYVFMSENPGMLWILGCHNSDFRNRG 720  
| | | | |  
Db 661 igatdlisvffsgytrkhwedyedtltlppsgetyvmsepgmlwlgchnsdfrnrg 720



```

QY 721 MTAALIKVSSCDKNAGDYEDSYEDISAVLLSKNNALIEPRSFSONSRHPSTROKOFNATPP 780
    |||||||
Db 721 mtaalikvsscdknagdyedsyedisavllsknnaleprsfsonsrhpstrokofnatpp 766
QY 781 VLKRHRQRETRRTTLOSDOEIDYDPTISVEMKKEDEFDIYDEDNOSPRSPQKTRRHYTA 840
    |||||||
Db 767 vlkrhrqrettrrttlosdoeidydptisvemkkedfdiydednqsprsqktrrhyta 826
QY 841 AVERLMDYGMSSSPHYLRNRAQSGSVQPKKVVQFEFTGSGFTQPLYRGELNHLGLGP 900
    |||||||
Db 827 averlmdygmsssphylnrraqsgsvpqkkvvgfegtfgstfqplryrgelnhlglgp 886
QY 901 YIRAEVDNINWTFRNQASRPYSFSSLSISYEDDQGAEPKRNKVPNETKTYEMKVOH 960
    |||||||
Db 887 yiraevedninwtfnrnqasrpysslsisyeddgagapkrnkvpnetktyemkvoh 946
QY 961 HMAPPKDEFDCANAVFSVDLEKOVHSGLLIGPLVCHNTNLNPAAGROVTOGEFALFT 1020
    |||||||
Db 947 hmappkdefdcakawayfsvdlekdvhsqllgpllvchntlnpahgrqvtvgefalft 1006
QY 1021 IFDEKSWYFTENMERNCRAPCNIQMEDPTFEKENYRFHAINCYIMDTLPGLVMAODQRT 1080
    |||||||
Db 1007 ifdekswyftenmerncrapcnigmedptfekenryfhaingyimdtlpglvmaodqrtr 1066
QY 1081 WYLLSMGSENHISIHFSGHVFTVRKKEEYKMAALNLYGVGFETVEMLPSKAGIMVECL 1140
    |||||||
Db 1067 wyllsmgsnenhishfsgvhftvrkkeeykmaalnlvgvgefetvemlpskagimvecl 1126
QY 1141 IGEHHAQMSTFLVYSNCKQPTPLGASGHIRDFOITASGOVGMAPKTLARLHYSGIN 1200
    |||||||
Db 1127 igenhaqmstflvysnckqptplgmasghirdfoitasgovgomapktlarlhygsina 1186
QY 1201 WSTKEPFMIKVDLAPMLIHGIKTOGARQKFSLSYISQFIIMYSADGKRWQTYRGNSGT 1260
    |||||||
Db 1187 wstkepfmikvdlapmlihgiktoGARQKFSLSYISQFIIMYSADGKRWQTYRGNSGT 1246
QY 1261 TLMVFGVNDSSGICAHNFNPPIIARIYIRLHPHYSIRSTLRLMELMGCDLNSGSMPLGME 1320
    |||||||
Db 1247 tlmvfgvndssgikahnfnppliiariyirhlphysirstlrlmelmgcdlnscsmplgme 1306
QY 1321 SKAISDAITASSYFTNMFAWSPSKARLHLQGRSNAMRPVNNPKEMLOVDFQKTMKYT 1380
    |||||||
Db 1307 skaisdaqitassyftnmfatwspkarlhlqgrsnamrpvnnpkemlovdqkcmkyt 1366
QY 1381 GVTTOGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKVFGQNDOSFTPVVNSLDPP 1440
    |||||||
Db 1367 gvttogvkslltsmyvkeflissqdgqhwctllffngkvkvfgqngdsftpvnsldpp 1426
QY 1441 LTRYLRHPQSWHQAIALMEVLCGEADPLY 1471
    |||||||
Db 1427 ltrylrhpgswvhqialimevlgceadply 1457

```

```

PF 08-DEC-1999; 99WO-US29169.
XX
PR 10-DEC-1998; 98US-0209916.
XX
PA (FAR ) BAYER CORP.
PI Cho M, Chan SY, Kelsey W, Yee H;
XX
DR WPI: 2000-431311/37.
XX
PT Producing cells expressing a protein having factor VIII procoagulant
PT activity especially, human factor VIII in an industrial scale, involves
PT expressing a vector comprising a sequence coding for factor VIII in
PT human cells
XX
PS Claim 7, Fig 1, 27pp; English.
XX
CC Producing cells expressing a protein having factor VIII procoagulant
CC activity, comprises contacting the cells with a vector comprising a
CC selectable marker and a sequence coding for the protein having factor
CC VIII procoagulant activity operably linked to a promoter. The cells
CC are then selected and individual clones expressing high levels of the
CC protein are isolated from the selected cells. The cells produced by
CC the method are not only useful for producing protein having factor
CC VIII procoagulant activity but also for producing adenovirus and
CC adeno-associated virus strains for gene therapy. The advantage of
CC having cells producing protein with factor VIII procoagulant activity
CC is that factor VIII protein can be produced on an industrial scale
CC in the range of 2-4 pg/cell/day. Human Burkitt's lymphoma (HKB)
CC cells provide a protein-free production system to produce not only
CC B-domain deleted factor VIII but also other therapeutic proteins. The
CC vector used in the method preferably comprises B-domain deleted
CC factor VIII (BDD-FVIII), a transcriptional unit for BDD-FVIII and a
CC selectable marker, dihydrofolate reductase (dhfr). In addition, a
CC terminal repeat sequence from Epstein-Barr virus is inserted into the
CC vector to increase integration efficiency.
XX
SQ Sequence 1438 AA:

```

Query Match 97.5%; Score 7674; DB 21; Length 1438;  
 Best Local Similarity 99.0%; Pred. No. 0;  
 Matches 1438; Conservative 0; Mismatches 0; Indels 14; Gaps 1;

```

QY 20 ATRRYLAGAVELSWDMQSDLGELPYDARPPRPVPSKPEFNTSVYKTLFVEFTVHLFN 79
    |||||||
Db 1 atrrylagavelswdmqsdlgelpydarpprvpskpfntsvyktllfveftvhlfn 60
QY 80 IAKRRPPMGLGPTIQAEVYDTVVYITLKNMASHYSLAHAVSVYWKASEGAEYDDQSQ 139
    |||||||
Db 61 iakrrppmngllgptlqaevydtvyltlknashpvsilhavsvywkasegaeyddqtsq 120
QY 140 REKEDKVPFGSGHYVWQVLKENGPMASDPLCTFYSTYSHVDLYKDNLSGILGALVCR 199
    |||||||
Db 121 rekedkvfpvgshyvwylkengpmasdpclctfysyshvdlykdnlsgilgallvcr 180
QY 200 EGSIAKEKTQTLHKFILLPAVEDEGKSMHSETKNSLMDRDAASARAMPKMHYGVYNR 259
    |||||||
Db 181 egslakektqltkflllfaivedegkswshsetknsldmgrdaasarawpkmhltvngyr 240
QY 260 SLPLGLGCHRRKSYWVWVIGMGTPPEVHSTFLGCHFTLVANHHQALESPTFLTAQTL 319
    |||||||
Db 241 slpllgchhrrksywvhwvigmgtpevhstflgchftlvnhrqaselsptfltaqtl 300
QY 320 MDLGGFLFCCHTSSHODMEAVYKVDSCPEEPQLRMKNNEEAEEDYDDDLTDSMDVYRF 379
    |||||||
Db 301 mdlggflfchshshodmeavykvdsceepqlrmknnaeaedydddltdsmdvryf 360
QY 380 DDDNSPFTQIRSVAKRHKPTWVHYTAEEEDWDYAPLVADDBRSYKSYQLNNGPQRIG 439
    |||||||
Db 361 dddnspftqirsvakrhpktwvhytaeedwdyaplvadddrsyksyqlnngpqrig 420
QY 440 RRYKKVPRMAYIDETKTRDAIQHESGILGPLLYGEGVDTLLIIRKNQASRYNYIPIHGI 499

```



```

Db 421 rkykvtimaytdetefkcreaighesgllgpllygevgdtllllfkqaspryllypbgil 480
QY 500 TDVRLPLYSRLPKGVKHLKDPILLPELEFKIKWIVTVDEGPTKSDPRLTRYSFVWME 559
Db 481 tdvrlplysrlpkgvkhlkdpillpelefkikwivtvdegptksdprltrysfivme 540
QY 560 RDLASGLILGLICYSKESVDGRGNIMSDKRNVLIFSVDENRSMXYLLENIORLPNPAG 619
Db 541 rdlasgllgllicyskesvdgrgnimskdnvllfsydenrswylenqrlilnpag 600
QY 620 VOLEDPPEQASINMHSINGVYFDSIOLSVCLHEVAWYILSIGAQDFLSVFFSGYTFKH 679
Db 601 vloedppegasimhsingyvfdslqlsvclhevawylisigqdtlsvffsgytfkh 660
QY 680 KMVYEDTLTFPEFSGEYFVSMENPGLMILCCHNSDPFRNRMATLAKVSSCDKNTGDYE 739
Db 661 kmvyedtltlfpesgetvfsmenpgylwlgchnsdfrnrgmtallkvsckdntgdye 720
QY 740 DSYEDISAYILSKNNATEPRSFSONSRHPTROKQFNATPPVLRHOREITRTLQSDOE 799
Db 721 dsyedisayilsknnaleprsfsonsrhptroqfnatppvlrhoraitrtlqsdqe 766
QY 800 EIDYDDITISVEMKKEDEPDYDEDENSPFSFOKTRHYFIAVERLMDYGMSSPHVLRN 859
Db 767 eidydditsvemkkedepdydedenspfsfokttrhyfiaverlmdygmssphvlrn 826
QY 860 RAQSGVSPQFKVAVFOETDGSFTQPLYRGELNHLGLGFIYRAVEDNIMVTFRNQAS 919
Db 827 raqsgsvpqfkvavfoetdgsftqplyrgelnhlglgfiyraavednimvtfrnqas 886
QY 920 RPYEYSSILSYEDDROGAPERRNFVKNPTKTYFMKVQHMMPTDEPFCKMAAFSD 979
Db 887 rpyeyssilsyeddrogaperrnfvknptktyfmkvqhmmptdepfckmaafsd 946
QY 980 VDLEKDVHSGILGILVCHTNTLPAHGRQVTVQEFALFTIPDETSMWYTEMNERCR 1039
Db 947 vdlekdvhsgilgvlvchtntlpanhgrqvtvqefalftipdetkswytemnercr 1006
QY 1040 APCNIOMEDPFENRYRPHAINGYIMDTLPGLVNAODRIRWYLLMSGNSNHSIHESG 1099
Db 1007 apcnlomedpfkenryrphaingyimdtlpglvnaodrirwyllmsgnsnhsihesg 1066
QY 1100 HVFVRKKEEKMLNLYPCGVFEFVEMLPKAGIWRVRECLIGHLHAGMSTFLVYSNK 1159
Db 1067 hvfvrkkeekmallypcgvfeftvmlpskagiwrveclighlhagmstflvysnk 1126
QY 1160 CQTPUGMASGHIRDFOITASGOYQMAPKLARLHYSGSINAMSTKEPFWIKVDLAPMI 1219
Db 1127 cqtpugmasghirdfoitassgoymapklarlhysgsinamstkepfwikvdlapmi 1186
QY 1220 IHGIRKTQGAROKSFSSLIISOFIMYSLDGKKWQTYRGNSGTGLWFRGNVDSGIRHNF 1279
Db 1187 ihgirktagaroksfssliisofimysldgkkwqtyrgnstgclwfrgnvdsigirhnlf 1246
QY 1280 NPPIIARIIRLHPHYRSRSLRMEILMGCDLNSCSMPRGMSKASISAQITASSYFTMMF 1339
Db 1247 nppliaarirhphyrslrslrmeilmgcdlnscsmprgmkskasisaqitassyftmmf 1306
QY 1340 ATWSPSKARLHLQGRSNAMRPQVNNPKEMLOVDFOKTKWVGTGTQGVKSLTSMYVEF 1399
Db 1307 atwspskarhlhlggrsnamrpqvnnpkewlqvdfoktkwvgtgtgvkslltssmykef 1366
QY 1400 LISSQDCHQWMTLFFONGKAVVFOGNODSTPPVNSLDPLLTRYLRIHPOSWVHQAIALR 1459
Db 1367 liissqdgqhwmtlffngkavvfognodstppvnsldplltrylrihposwvhqialr 1426
QY 1460 MEVIGCEAODILY 1471
Db 1427 mevlgceagdily 1438

```

RESULT 6

```

AAR12971
ID AAR12971 standard; protein; 1440 AA.
XX
AC AAR12971;
XX
DT 02-OCT-1991 (first entry)
XX
DE Factor VIII:SQ.
XX
KW Factor VIII; B domain; haemophilia.
XX
PN MO9109122-A.
XX
PD 27-JUN-1991.
XX
PF 06-DEC-1990; 90WO-SE00809.
XX
PR 15-DEC-1989; 89SE-0004239.
XX
PA (KABI ) KABIVITRUM AB.
XX
PI Almstedt AB, Hellstrom EM, Larsson K, Lind P, Sandberg HI, Spira J;
PI Sydow-Backman M;
XX
DR WP1; 1991-208148/28.
XX
PT Recombinant human factor VIII deriv. deoxyribonucleic acid -
PT encoding protein comprising two chains linked by segment of B domain.
XX
PS Disclosure; Fig 1; 35pp; English.
XX
CC The protein is a fusion between Phe 742 and Ser 1637 of the factor
CC VIII protein (factor VIII:SQ). In order to produce a
CC factor VIII deletion derivative that can be produced in vivo and/or
CC in vitro, to a two chain protein consisting of polypeptide chains of
CC 90 kd and 80 kd, the amino acid sequences surrounding Arg 740 and
CC Arg 1648 have to be conserved in order to preserve the structural
CC requirements for correct cleavage. In this example, amino acids 743
CC to 1636 of the full-length factor VIII polypeptide are deleted. A
CC new polypeptide chain is obtd. where there are 14 amino acids
CC linking Arg 740 and Arg 1648. Of these 14 amino acids, the sequence
CC of the five N-terminal ones directly corresponds to the five amino
CC acids following Arg 740 in full-length factor VIII. Also, the sequence
CC of the 12 C-terminal amino acids of the above 14 amino acids fragment
CC directly corresponds to the 12 amino acids preceding Glu 1689 in full-
CC length factor VIII, thus creating a 3 amino acid overlap between the
CC N- and C-terminal regions of the B-domain.
CC The factor VIII deriv. is useful for treating haemophilia or
CC haemophilia A. It has the biological characteristics of plasma derived
CC factor VIII.
CC In order to index this example, the factor VIII:QD amino acid
CC sequence was retrieved from W08800831 (AAB80265).
CC The amino acid numbering in the above comments is reproduced from the
CC fig. description in the specification. Note that Arg 740 is Arg 742
CC in the indexed sequence, etc. Also, Asn 745 (N-terminal link overlap)
CC is Asp 747 in AAB80265, but indexed as Asn to reproduce the fusion
CC fragment as shown in fig.1, and Asn 745 (C-terminal link overlap).
XX
SQ Sequence 1440 AA:

```

Query Match 97.1%; Score 7642.5; DB 12; Length 1440;  
 Best Local Similarity 98.7%; Pred. No. 0;  
 Matches 1434; Conservative 1; Mismatches 3; Indels 15; Gaps 2;

```

QY 20 ATRRYVLGAVELSDWYMSDGLGPVDARPPRPVKSPPFWTSVYKKTLFVEFTHLFN 79
Db 2 atrryylgavelsdwymdsdglgpvdarpprpvksppfwtsvykktlfveftdhfn 61
QY 80 IAKRPPEWMLGPTIOAEVYDVTVITLKNASHPVSLHAGVSYWKASEGAEYDDQTSQ 139
Db 62 iakrppewmlgptioaevydvvtvitlknashpvsilhavgsywkasegaeyddqtsq 121

```



```

QY 140 REKEDKVPGGSHYVWQVLKENGPMASDPLCLTYSTLSHVDLVKDLNSGLIGALLVCR 199
    |||
Db 122 rekeddkvfpgshlywvwlkengpmasdpclctysylshvdlvkdlnsgllgalllvc 181
QY 200 EGSJLAKKKTQTLHKFTLLFAVFEDEGSMHSERTKNSLMQORDAASARAWKMHNTVYVR 259
    |||
Db 182 egslakekethlkhkfillfavidegskwshseknsimgdrdaasarawkmbhvtvgyvr 241
QY 260 S-LPLGLICHRKSVYMHVIGMCTPEVHISFLEGHTFLVRNHRQASLESPITFLAQT 318
    |||
Db 242 slpplglchtrksywhvlgmgtpevhslfleghtflvrnhqaslelspitfltaql 301
QY 319 LMDLQFLFLCHISSHQHDGMEAYKVDSCPEEPOLRMKNNEBAEDYDDLTDSEMDVVR 378
    |||
Db 302 lmdlqgfillfchisshqhdgmeaykvdscpeepqlrmkneeaedyddldtsemdlvvr 361
QY 379 FDDQSPSFIOIRSVAKKHPTVHVYIAEEDMDVAPVLAAPDDPSYQYLNNGPQRT 438
    |||
Db 362 fdddspsfliqirsvakkhptcvhylaeeedwdyaplvlapddtsyksqylngpqrl 421
QY 439 GRKYKRVFMAVTDDEFTKREALIOHESGILGPLYGEVDTLLITFRNQASRPYNIYPHG 498
    |||
Db 422 grkykvrvmaytdetfktrealiohesgllgpllygevgdtlllftknqasrpynllyphg 481
QY 499 ITDVRPLYSRLPKGVKHLKDFPILPGEIEKYKWTYVEDEGPTKSDPCLTRYSSFVNM 558
    |||
Db 482 ltdvrplysrllpkgvkhlkdfpilpgeifkykwtyvedegptksdprcltryssfvnm 541
QY 559 ERDLASSGILGPLICRKESVDORGNQIMSDKRNVLFSFEDENRSYLTENQRIPLPNA 618
    |||
Db 542 erdlasgllgpllicykeavdgrgnqmsdkrnvllfsyidenrsytleenqrlfplpna 601
QY 619 GVOLEDPFEOASNIMHSINGYFDSLQSLVCLHEVAWYILSIGAOTDLSVFFSGYTPK 678
    |||
Db 602 gvqledpfqasnlnhsingyfdslqslvclhevawylisigqtdflsvffsfytpk 661
QY 679 HKMYEDTLTLPFSGEYVFMSENPGLWILCNSDFRNKGMTALLKVSCKDKNTGDY 738
    |||
Db 662 hkmyedtltpfsgetvfmsempglwlilgchnsdfnrgmtallkvsckdkntgdy 721
QY 739 EDSYEDISAVYLSKNNAITPRFSQNSRHRPSTROKOFNTPRVLYKHOREIRTTQSQO 798
    |||
Db 722 edsyedisaayllsknnaileprsfsgn-----pvlkrhngelctcltqsdq 767
QY 799 EEIYDDTISYEMKKEDFDIYDEENQSPRSFOKTRHYFIAVERLAMYGMSSPHVLR 858
    |||
Db 768 eeiyyddtisvemkkedfdiyedengsprsfqktrhyfiaaverlwygmssphvlr 827
QY 859 NRAQSGSVPOKRVVFOERTDSFTQPLYRGLNHLGLGPLYIRAEVEDNIMVTERNOA 918
    |||
Db 828 nraqsgsvpqfkkvfygeftdgsftqplyrgehlglgplyiraevednimvternga 887
QY 919 SRPYSFYSLSIYEDDOGAAPRKPNFKPNETKTYFMKVOHNMAPTQKPEPCKAMAYS 978
    |||
Db 888 srpysfysllsyeedqggaeptrkpnktyfkvnqhmarpckelockaamays 947
QY 979 DVDEKDVHSGILGPLYCHNTTLNPAHGROYTVOEFALFPIFDETSKMYTFENMERNC 1038
    |||
Db 948 dvdekdvhsgllgpllychntlnpangroytvtgefalfpifdetksmytfenmernc 1007
QY 1039 RAPNCIQMEDPTFKENYFPAHNGITMDTLPGLVMAODORIMWYLSMGSNENHSIHS 1098
    |||
Db 1008 rapnciqmedptfkennyfphaingyimdtlpglvmaodqrlwylmsnensihshs 1067
QY 1099 GHVFTVRKKEVKMALVNLVPGVEFTEMLPSKAGTWRECLIGELHAGMSTFLVYNS 1158
    |||
Db 1068 ghvftvrkkeekmalvnlvpgvetemlpskagtwrecligenlhaqgmstflvyns 1127
QY 1159 KCQTPPLGASGHIRDFQTASGOYCOMAPKLARLHYSGSINAMSTKEPSSWKVDLAAM 1218
    |||
Db 1128 kcpplplgmasghirdftqtaasgyqwapklarlhysgsinamstkepswkvldlaam 1187
QY 1219 ITHGKTGAGAROKFSLYISQFIWYSLDGKKWQTYRGNSTGTLMVFGNVDSGIRKNI 1278

```

```

Db 1188 |||
    |||
QY 1279 ENPPIARIYRIHPHYSTRSLRMELMGCDLNSCSMPLGMSKASIAOAITASYPNM 1338
    |||
Db 1248 fnpplariyrlhphystlrlmelmgcdlnscsmplgmeskaiadaqlaasyfntm 1307
QY 1339 FATWSPSKARLHLQGRSNMARPQVNNPKEMLOVDQKTKVYGVTTQGVKSILTSMYKE 1398
    |||
Db 1308 fatwspskarlhlqgrsnmrpqvnnpkemlvdqtkmkvgtvtgvtksiltsmyke 1367
QY 1399 FLISSQDGHQWTLFQNGKRVFQGNODSFMPVNSLDPPLTRYLRNIHPQSWVHQIAL 1458
    |||
Db 1368 flissqdgqhwtlffqngkrvfvgngdsfmpvnsldppltrylrnhpqswhqial 1427
QY 1459 RMEVLGCEADQLY 1471
    |||
Db 1428 rmevlgceaqdly 1440

RESULT 7
AAW18670
ID AAW18670 standard; Protein; 1661 AA.
XX
AC AAW18670;
XX
DT 10-AUG-1997 (first entry)
XX
DE Factor VIII-dB695-HCII.
XX
KW Factor VIII-dB695-HCII; heparin cofactor II; blood coagulation;
KW blood clotting; procoagulant; anticoagulant; antithrombotic;
KW haemophilia; gene therapy.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Domain 731..760
FT /label= HCII
FT /note= "heparin cofactor II acidic domain"
XX
PN WO9118315-A1.
XX
PD 22-MAY-1997.
XX
PE 13-NOV-1996; 96MO-EP04977.
XX
PR 13-NOV-1995; 95US-0558107.
XX
PA (IMMO ) IMMUNO AG.
XX
PI Voorberg JJ.
XX
DR WPI: 1997-289291/26.
DR N-PSDB: AAT69811.
XX
PT Hybrid Factor VIII with modified activity, comprises region from
PT donor anticoagulant or antithrombotic protein - useful for treatment
PT of coagulation disorders
XX
PS Claim 11; Page 52-60; 96pp; English.
XX
CC Factor VIII-dB695-HCII (AAW18670) is a hybrid protein in which amino
CC acids 712-736 of Factor-dB695 (Factor VIII del868-1562) B-domain
CC are replaced by amino acids 51-80 from the acidic region (and
CC potential thrombin-binding site) of human heparin cofactor II
CC (HCII). It is the expression product of Factor VIII-dB695-HCII
CC DNA (AAT69811) in plasmid pCLB-dB695-HCII. The hybrid protein, which
CC can be expressed using gene therapy techniques, has increased
CC procoagulant activity owing to the HCII acidic region, and can be
CC used to treat blood coagulation disorders such as haemophilia A.
XX
SQ Sequence 1661 AA;

```



Query Match 96.8%: Score 7616: DB 18: Length 1661:  
 Best Local Similarity 87.0%: Pred. No. 0:  
 Matches 1450: Conservative 5: Mismatches 10: Indels 202: Gaps 4:

QY 1 MQEILSTCFELCLLRFCFSATRRYYLGAVELSWDMQSDGLCELPVDAFRPVRKSEPFN 60  
 1 meielstcfelcllrfcfsatrryylgavelswdymqsdglcelpvdarfrprvrksfpfn 60

QY 61 TSVYKKTLPVEFTVHLFNIAKRRPPMGLGPTIOAEVDTVTYITLKNMASHVSLHAY 120  
 61 tsvykktlpveftvhlfniakrppmglgptioaevdtvttyitlknmashvslhay 120

QY 121 GVSYSKASEGAEYDDQTSOREKEDKVPGGSHYVQVLEKNGPMASDPCLCTYSLSH 180  
 121 gvsyskasegaeyddqtsorekedkvfpgshyvwqvlengkpmasdpclctyslsh 180

QY 121 GVSYSKASEGAEYDDQTSOREKEDKVPGGSHYVQVLEKNGPMASDPCLCTYSLSH 180  
 121 gvsyskasegaeyddqtsorekedkvfpgshyvwqvlengkpmasdpclctyslsh 180

QY 181 VDLVKDLNSGLIGALVCREGSLAKEKTQTLHKEFILLFAVEDEGKSMHSETKNSLMDRD 240  
 181 vdlvkdlnsgligalvcregslakektqtlhkefillfavedegksmhsetknsldmrd 240

QY 181 VDLVKDLNSGLIGALVCREGSLAKEKTQTLHKEFILLFAVEDEGKSMHSETKNSLMDRD 240  
 181 vdlvkdlnsgligalvcregslakektqtlhkefillfavedegksmhsetknsldmrd 240

QY 241 AASARAPKMHNTYNGYVNRSLPGLIGCHRSVYWHVIGMCTTPEVHSIFLEGHTFLVRNH 300  
 241 aasarapkmhntvngyvnrslpgligchrsvywhvigmcttpevhsifleghtflvrnh 300

QY 241 AASARAPKMHNTYNGYVNRSLPGLIGCHRSVYWHVIGMCTTPEVHSIFLEGHTFLVRNH 300  
 241 aasarapkmhntvngyvnrslpgligchrsvywhvigmcttpevhsifleghtflvrnh 300

QY 301 RQASLETSPITTEFLAOTLLMDLGFLFCHTSSHQDHGMEXYKVVDSCEPPOLRMKNE 360  
 301 rqaletspitteflaotllmdlgflfchttsshqdhgmexykvvdscpeppolrmkne 360

QY 301 RQASLETSPITTEFLAOTLLMDLGFLFCHTSSHQDHGMEXYKVVDSCEPPOLRMKNE 360  
 301 rqaletspitteflaotllmdlgflfchttsshqdhgmexykvvdscpeppolrmkne 360

QY 361 EAEDYDDDLTDSEMDVYRFDDDNPSFQIQRVAKKPKTWHVIAAEEDMDAPVLA 420  
 361 eaedydddltdsemdvyrfdndnpsfqiqrvakpkptwhviaaeedmdapvla 420

QY 361 EAEDYDDDLTDSEMDVYRFDDDNPSFQIQRVAKKPKTWHVIAAEEDMDAPVLA 420  
 361 eaedydddltdsemdvyrfdndnpsfqiqrvakpkptwhviaaeedmdapvla 420

QY 421 PDBRSYSQTLNNGPQIRGRKKRVPMAYTDETKREAIQHEGSLGPLLYGEVDTL 480  
 421 pddrsysqtlngpqrigrkkrvpmaytdektreaiqhegslgplylgevgdtl 480

QY 421 PDBRSYSQTLNNGPQIRGRKKRVPMAYTDETKREAIQHEGSLGPLLYGEVDTL 480  
 421 pddrsysqtlngpqrigrkkrvpmaytdektreaiqhegslgplylgevgdtl 480

QY 481 LIIFKNASRPYNTYPHGINDVRLYRRLPKGVKHLKDPILPGLFVKTKWTYVDEGP 540  
 481 liifknasrpyntyphgindvrllyrllpkgvkhlkdpilpglfvtkwtvydegp 540

QY 481 LIIFKNASRPYNTYPHGINDVRLYRRLPKGVKHLKDPILPGLFVKTKWTYVDEGP 540  
 481 liifknasrpyntyphgindvrllyrllpkgvkhlkdpilpglfvtkwtvydegp 540

QY 541 TKSDPRCLTRYSSFVMMERDLASGLIGPLLYCYKESVDQNGOIMSDKRNVLFSVDE 600  
 541 tksdprcltryssfvmmerdlasgligplylcykesvdqngoisdkrnvlfsvde 600

QY 541 TKSDPRCLTRYSSFVMMERDLASGLIGPLLYCYKESVDQNGOIMSDKRNVLFSVDE 600  
 541 tksdprcltryssfvmmerdlasgligplylcykesvdqngoisdkrnvlfsvde 600

QY 601 NRSWYLTENIQRLPNPAGVQLEDPEFOASNIHSHINGYVDSIQLSVCLHEVAYWYLS 660  
 601 nrswylteniqrlpnpagvqledpefoasnishshingyvdsiqlsvclhevaywyls 660

QY 601 NRSWYLTENIQRLPNPAGVQLEDPEFOASNIHSHINGYVDSIQLSVCLHEVAYWYLS 660  
 601 nrswylteniqrlpnpagvqledpefoasnishshingyvdsiqlsvclhevaywyls 660

QY 661 IGAOTDLSVFEFGYTPFKHKVYEDTLTFPFGSEYFMENENGLMILGCHNSDFRNRG 720  
 661 igatdlsvfefgytpfkhkvyedtltfpfgsetyfmennlglmllgchnsdfrnrg 720

QY 661 IGAOTDLSVFEFGYTPFKHKVYEDTLTFPFGSEYFMENENGLMILGCHNSDFRNRG 720  
 661 igatdlsvfefgytpfkhkvyedtltfpfgsetyfmennlglmllgchnsdfrnrg 720

QY 721 MTALLKTVSSC---DKNTGDY-----BDSYEDIASYLLSKNNATPEPSSONSRRPS 769  
 721 mtaallkvtssc---dkntgdy-----bdsyediasyllsknnatpexpssonsrrps 769

QY 721 MTALLKTVSSC---DKNTGDY-----BDSYEDIASYLLSKNNATPEPSSONSRRPS 769  
 721 mtaallkvtssc---dkntgdy-----bdsyediasyllsknnatpexpssonsrrps 769

QY 770 TROKOFNAT----- 778  
 770 trokofnat----- 778

QY 775 TRGQFNATCIPENDIEKTDQPFHRTCPMKIQNVSSDILMLLRGPTPHGLSLSLDGE 834  
 775 trgqfnatcipendiektdqpfhrtcmpkignvssdilmllrgptphglslsldge 834

QY 775 TRGQFNATCIPENDIEKTDQPFHRTCPMKIQNVSSDILMLLRGPTPHGLSLSLDGE 834  
 775 trgqfnatcipendiektdqpfhrtcmpkignvssdilmllrgptphglslsldge 834

QY 835 AKYETSDSPSPAIIDNSLSEMTHRPQLHNSGDMVFPRESGLRLNEKJLGTADPL 894  
 835 akyetfddspspaidnslsemthrpqlhnsghdmvfpresglrlneklgtadpl 894

QY 835 AKYETSDSPSPAIIDNSLSEMTHRPQLHNSGDMVFPRESGLRLNEKJLGTADPL 894  
 835 akyetfddspspaidnslsemthrpqlhnsghdmvfpresglrlneklgtadpl 894

QY 895 AWDNHYGTQJPKKEWKSQEKSPKTAFKKDDLLSLNACSNHAIASINQGNKPELEVT 954  
 895 awdnhygtqjpkkewksqekspktafkddllslnacsnhaiaainqgnkpelevt 954

QY 895 AWDNHYGTQJPKKEWKSQEKSPKTAFKKDDLLSLNACSNHAIASINQGNKPELEVT 954  
 895 awdnhygtqjpkkewksqekspktafkddllslnacsnhaiaainqgnkpelevt 954

QY 779 -----PVLAKRHQREITRTTLOSDOEIDYDDTJSVEKKKEDFDIYDEDN 824  
 779 -----pvlakrhqreitrttlosdoeidyddtjsvekkkedefdiydedn 824

QY 779 -----PVLAKRHQREITRTTLOSDOEIDYDDTJSVEKKKEDFDIYDEDN 824  
 779 -----pvlakrhqreitrttlosdoeidyddtjsvekkkedefdiydedn 824

QY 955 WAKGRTERLCSPGPVLRKHGRETITCLTQSGDEEIDYDDTJSVEKKKEDFDIYDEDN 1014  
 955 wakgarterlcsppvllrkhgreitrtcltqsgdeeidyddtjsvekkkedefdiydedn 1014

QY 825 QSPRSFOKTRHRYFLAAVERLMDYGMSSSPHYLRNRAQSGSVPOFKKVFQETDGSFTQ 884  
 825 qsprsfoktrhryflaaverlmdygmsssphylnraaqsgsvpofkkvfqetdgsftq 884

QY 1015 QSPRSFQKTLRHFLAAVERLMDYGMSSSPHYLRNRAQSGSVPOFKKVFQETDGSFTQ 1074  
 1015 qsprsfqktlrhflaaverlmdygmsssphylnraaqsgsvpofkkvfqetdgsftq 1074

QY 885 PLYRGELNEHLGLLPYIRAEVEDNINWTFRNQASRPYSFYSLLSYEEDROGAEPKRN 944  
 885 plyrgelnehlglp yiraevedninwtf rdnqasrpy sfysllsyeedrogaepkrn 944

QY 1075 PLYRGELNEHLGLLPYIRAEVEDNINWTFRNQASRPYSFYSLLSYEEDROGAEPKRN 1134  
 1075 plyrgelnehlglp yiraevedninwtf rdnqasrpy sfysllsyeedrogaepkrn 1134

QY 945 FVKPNEKTYRWKQVHNAAPRKDEFCKAAMVFSVDLEKRVHSGLLGPLVCHTNTLP 1004  
 945 fvkpnektyrwkqhnaaprkddefckaa mvfsvdlekrvhsgllgplvchtntlp 1004

QY 1135 FVKPNEKTYRWKQVHNAAPRKDEFCKAAMVFSVDLEKRVHSGLLGPLVCHTNTLP 1194  
 1135 fvkpnektyrwkqhnaaprkddefckaa mvfsvdlekrvhsgllgplvchtntlp 1194

QY 1005 AHGRQVTVQEEALFFETJFDEKSWYFTEENMRNCRAPCNIMEPTPEKFNRFHAINGY 1064  
 1005 ahgrqv tvqealffetj fdekswy fteennrcrapcnimeptpekfnrfhaingy 1064

QY 1195 AHGRQVTVQEEALFFETJFDEKSWYFTEENMRNCRAPCNIMEPTPEKFNRFHAINGY 1254  
 1195 ahgrqv tvqealffetj fdekswy fteennrcrapcnimeptpekfnrfhaingy 1254

QY 1065 MDTLPLGLVMAODORIHWYLSMGSNENIHSIHFGHVFYVKKKEKMYALNYLPVFEET 1124  
 1065 mdlplglvmaodori hwy lsmgsnenihs ihfghvfvykkkemyalnylpvfeet 1124

QY 1255 MDTLPLGLVMAODORIHWYLSMGSNENIHSIHFGHVFYVKKKEKMYALNYLPVFEET 1314  
 1255 mdlplglvmaodori hwy lsmgsnenihs ihfghvfvykkkemyalnylpvfeet 1314

QY 1125 VEMLPKSKAGIWRVCLIGEHHLAAGMSTFLVYSNKCOTPLGASGHIRDPOITASGORYQ 1184  
 1125 vemlpkskagiwrvc ligehlhaagmstflvysnkcotplgasghirdpoit asgoryq 1184

QY 1315 VEMLPKSKAGIWRVCLIGEHHLAAGMSTFLVYSNKCOTPLGASGHIRDPOITASGORYQ 1374  
 1315 vemlpkskagiwrvc ligehlhaagmstflvysnkcotplgasghirdpoit asgoryq 1374

QY 1185 WAPKLARLHYSGSTINAMSTKEPFSIKVLDLAPMIIHGIRKQGRKRFSSLYISQFTIMY 1244  
 1185 wapl arlhysgstinamstkepf sikvldlapmiihgir kqgrkrfsslyisqftimy 1244

QY 1375 WAPKLARLHYSGSTINAMSTKEPFSIKVLDLAPMIIHGIRKQGRKRFSSLYISQFTIMY 1434  
 1375 wapl arlhysgstinamstkepf sikvldlapmiihgir kqgrkrfsslyisqftimy 1434

QY 1245 SLDGKKNQOTYRGNSTGTLMVFGNVDSGIGKHNIFNPPIIRYTRLPHTHYSISTLME 1304  
 1245 sldgkknqotyrgnstgtlmvfgnvds gikhnifnp ppiir ytrlphthysistlme 1304

QY 1435 SLDGKKNQOTYRGNSTGTLMVFGNVDSGIGKHNIFNPPIIRYTRLPHTHYSISTLME 1494  
 1435 sldgkknqotyrgnstgtlmvfgnvds gikhnifnp ppiir ytrlphthysistlme 1494

QY 1305 LMGCDLNSCKPCLMESKASISDAQITASSYTNMFATWSPSKARLHLOGRSNANRPVYN 1364  
 1305 lmgcdl nsckp clmeskas isdaq itass ytnmf atwsp skarl hlogrsnanrpvyn 1364

QY 1495 LMGCDLNSCKPCLMESKASISDAQITASSYTNMFATWSPSKARLHLOGRSNANRPVYN 1554  
 1495 lmgcdl nsckp clmeskas isdaq itass ytnmf atwsp skarl hlogrsnanrpvyn 1554

QY 1365 PKEWLOYDFOKTMTKVTGTTQGVKSLTSMVKEFLTSSQDGHQWTFPONGKVKYEOG 1424  
 1365 pkewloydfoktmtkvtgttqgvkslt smvkefl tssqdghqwtfpongkvkyeog 1424

QY 1555 PKEWLOYDFOKTMTKVTGTTQGVKSLTSMVKEFLTSSQDGHQWTFPONGKVKYEOG 1614  
 1555 pkewloydfoktmtkvtgttqgvkslt smvkefl tssqdghqwtfpongkvkyeog 1614

QY 1425 NQDSFTPVNSLDPLRLTRYLRIPQSNVHQAIRMEVLCGEADLY 1471  
 1425 nqdsftpvnsldplrltrylr ipqsnvhqairmev lcgeadly 1471

QY 1615 NQDSFTPVNSLDPLRLTRYLRIPQSNVHQAIRMEVLCGEADLY 1661  
 1615 nqdsftpvnsldplrltrylr ipqsnvhqairmev lcgeadly 1661

RESULT 8  
 AAP80265  
 ID AAP80265 standard; protein; 1516 AA.  
 AC AAP80265;  
 AC AAP80265;  
 DT 10-OCT-1990 (first entry)  
 XX  
 DE Modified factor VIII:C sequence with the Q744-D1563 deletion.  
 DE Modified factor VIII:C; maturation polypeptide; haemophilia;  
 KW blood coagulation; QD deletion.  
 KW  
 OS Homo sapiens.  
 OS  
 PN W080800831-A.  
 PN  
 PD 11-FEB-1988.  
 PD  
 XX 31-JUL-1987; 87WO-US01814.  
 XX  
 PF 31-JUL-1987; 87WO-US01814.  
 PF  
 PR 01-AUG-1986; 86US-0893375.  
 PR  
 XX (BIOU ) BIOGEN NV (PASE//).  
 XX  
 PA PASEK MP;  
 PA  
 PI  
 XX



DR WPI, 1988-049866/07.  
DR N-PSDB: AAN80444.  
XX New DNA sequences encoding modified factor VIII:C - with deletion of DNA  
PT encoding maturation polypeptide, useful for high yield transformation.  
XX  
PS Claim 3; Page 51-52-53-54; 97pp; English.  
XX  
CC A major part of the sequence encoding the maturation polypeptide of  
CC factor VIII:C is deleted, i.e. Gln 744 - Asp 1563. The 6D deletion  
CC retains approximately 90 amino acids of the maturation polypeptide  
CC (four amino acids at the N-terminal end and 86 amino acids at  
CC the C-terminal end). The full length Factor VIII:C cDNA has two  
CC changes with respect to the published sequence (EPO application 160457):  
CC CTG to CTA at Leu 242 and TTC to CTC change at amino acid residue 1880  
CC (Phe to Leu).  
CC The product is produced in approx. 20 times higher  
CC yields than previous recombinant produced factor VIII:C and are more  
CC easily purified. The peptide is used for treating haemophilia A, both  
CC acute and prolonged bleeding.  
CC See also AAN80446 and AAN80447.  
XX  
XX Sequence 1516 AA:  
Query Match 96.8%; Score 7615.5; DB 9; Length 1516;  
Best Local Similarity 94.8%; Pred. No. 0;  
Matches 1436; Conservative 5; Mismatches 11; Indels 63; Gaps 3;  
QY 20 ATRRYIYIGAVLSMDYMSDGLGELPYDARFPPRVKSPFPNTSVYKKTLFPEFVHLEN 79  
DB 2 atrryyigavelsmdymdsdglgelpydarfprrvpksfpntsvykkltfpevfchlfn 61  
QY 80 IAKPRPMGGLGPTIAEYVDTVITLKNMASHPVSLHAGVSYWKASEGAEVDQTSQ 139  
DB 62 iakprppmgjllgptigevydvttvltkmasbpvslhavgyswkasegeydddcsq 121  
QY 140 REKEDKVFPGSGHTYVMQVLKENGPMASDPLCLTYSLSHVDLVKLSLIGALLVCR 199  
DB 122 rekeddkvfpgsghtyvmqvllkengpmasdpcltysylshvdlvklslngllvcr 181  
QY 200 EGSIAKEKTOYLKFIILFAVFEDEKSWHSEKNSLMODRDAASARAPKMTTVGVYNR 259  
DB 182 egslakektolhkfllfayfedeqskwhseksnlmodrdaasarpkmttvnyvr 241  
QY 260 S-LPGLIGCHRSYVWHVIGMGTREPVHSIFLEGHTFLVRNHRQASLESPITFLAQL 318  
DB 242 slppllgchrsyvwvhvigmgtrepvhsifleghtflvrnhrqasleisplftlaql 301  
QY 319 LMDIGQFLLECHISSHQHDGMEAVYKVDSCPEEPQLRMKNNEADYDDDLTDEMDVVR 378  
DB 302 lmdigqflfchishsqhdgmeaykvkvdscpeepqlrmkneeadydddltdsemdlvr 361  
QY 379 FDDDNSSSFIOIRSVAKKHRTKTVNHYIAAEEDMDYARLVLPDRRSKYSQSLNNGPQRI 438  
DB 362 fdddnsssfioirsvakkhrtkvtvnyiaaeedmdyarlvpdrsrkysqslnngpqri 421  
QY 439 GRKKKRYFMAVYDTEFTRTAIOHESGILPRLGEGVDTLLIFKNQASRPVNIYPHG 498  
DB 422 grkkkryfmayvdtetftrtaiohesgilprllgyevgdtlllffkngasrpvniyphg 481  
QY 499 ITDVRLYSRRLPKGVKHLKDFPILPGEIFRYKWTVTVEDEGFTKSDPCLTRYSSFVNM 558  
DB 482 itdvrlpysrrlpkgyvkhldfprilpgeifrykwtvtvedegftksdpcltryssfvm 541  
QY 559 ERDLASGLIGLLCYKESVNOQRNOIMSDKRNVLFSVPDENRKYVLTENIQRLPNPA 618  
DB 542 erdlasgligllcykesvnoqrnoimsdkrnvlfsvpdenrkyvlteniqrlpnpa 601  
QY 619 GVOLEDEFGASNMHSINGVVDLSVCLHVEAVYWTIISGQTFPLSFEFGYFVK 678  
DB 602 gvoledefgasnmhsingvvdlsqsvclhveawywtiisgqtdflsvfisygflvk 661

QY 679 HKMYEDTTLTLPFSGEFTPKSMENPGLMILGCHNSDFRNRMATALLKVSSCDKNQDGY 738  
DB 662 hkmyedttltlpfsgetfpmksmnpglmlilgchnsdfrnrmatallkvsscdkntgdy 721  
QY 739 EDSYEDISAYLKSKNALEPRFSQON-----SRHSTNQKPF----- 775  
DB 722 edsyedisaaylksknalesprfsqdlawdnhgytqlpkewksqespektafkkkt 781  
QY 776 -----NATPVLRKHOREIRRTLQS 796  
DB 782 lslsacsenhaiaaimgnkpkelevetwakgartericsqnpvylkrhgeirttlqgs 841  
QY 797 DQEEIDYDPTLSVEMKKEDFDYDEENOSPRSPQKTRHYFAAVERLMDYGMSSSPHV 856  
DB 842 dqeedidytlsvemkkedfdydedensprsfqktrhyfiaaverlwdygmsssphv 901  
QY 857 LNRNQSQSVQPKRVVQOEFTDGFQPLVYRGELNEHLGLGYIAAEVADNIMWFNRN 916  
DB 902 lnraqsgsvpqfkkrvvqeftdgsftqplvyrgelnehlglyiaavednwmvfrn 961  
QY 917 QASRPYSFSSLSIYEEDOROGAEPKRNFKVFNETHKTFMKVQHMMAPTKDEPCKAMAY 976  
DB 962 gasrpyssfslsiyeedqrgaepkrnfkvpnektkyfwkvqhmaptkdeickaway 1021  
QY 977 FSDVDLEKDVHSGILGPIVCHTNTLNPAHGRQVTVQGEFALFTIPDETCKSWYTFENMR 1036  
DB 1022 fsdvdlkdvhsgilgpllvchntlnpahgrqvtyvqgefalfitpdetckswytfenmr 1081  
QY 1037 NCRAPCNIOEDPTEPKENVRFPAINGYIMDLPGVYMAODORIMRYLLSMGNSNHSIH 1096  
DB 1082 ncrapcnlmedptckenvrfpaignyimdltpgvymaodorimryllsmgnsnhsih 1141  
QY 1097 FSGHVFYVKKKEEYKMAVLYNLYPGVEFVEMLPKAGITWRVECHLGEHLHAGMSTFLVY 1156  
DB 1142 fsghvfyvrkkeeykmalynlypgvfevemlpkagitwrvechlgehlhagmstflvy 1201  
QY 1157 SNKCOPTPLGASGHTRDQITASGOYGMARPLARLHSGSINAMWSKEPWSIKVULLA 1216  
DB 1202 snkcqptplgmasghtrdftqitassgygwapklarlhysgsinawskewfswikvulla 1261  
QY 1217 PMIHGIRKTOGAKROKESLSYISOPTIMYSLDGKKMQYRGSTGTLMWFPGNVDSGKIH 1276  
DB 1262 pmihgirktoarakrokslsyisoptimysldgkkmqyrgstgtlmwfgnvdsqkikh 1321  
QY 1277 NIFNPITARIYRIAPRTYISIRSTLMEIMGCDLNSGSMPLGMSKASISDAQTASSYFT 1336  
DB 1322 nifnpitariyrlaprtysirstlmeimgcdlnscmplgmskatsdaqtaassyft 1381  
QY 1337 NMFATWSPSKARLHLOGSNAMRPOVNNPKFEMLOYDFOKTMKVTGVTTCGYKSLITSMYV 1396  
DB 1382 nmfatwspskarlhlogsnamrpoynnnpkfemloydfoktmkvttgvttygykslitsmyv 1441  
QY 1397 KEFLTSSODGHOWLTFONKVKVFOGNODSFTPVNYSLLRPILLTRLRHPDSWVHOI 1456  
DB 1442 kefltssodgqhwcltfqngkvkvfgngqdsfcpvnyslldrpilltrlrhpswvhqi 1501  
QY 1457 ALRMEVLGCEAQDILY 1471  
DB 1502 alrmevlgceaqdily 1516  
RESULT 9  
AAB48842  
ID AAB48842 standard; protein; 1424 AA.  
XX  
AC AAB48842;  
XX  
DT 13-MAR-2001 (first entry)  
XX  
DE Mutant mature human factor VIII, SEQ ID NO:5.  
XX  
KW Factor VIII; human; B domain; LRP-mediated plasma clearance;  
KW receptor-dependent clearance; receptor-independent clearance;



KW half-life; haemophilia; mutant; mutein.  
XX  
OS Homo sapiens.  
XX  
FN W020071714-A2.  
XX  
PD 30-NOV-2000.  
XX  
PE 24-MAY-2000; 2000WO-US14111.  
XX  
PR 24-MAY-1999; 9905-0135847.  
XX  
PA (AMNA-) AMERICAN NAT RED CROSS.  
XX  
PI Saenko EL, Strickland DK;  
XX  
DR WPI; 2001-025163/03.  
XX  
PT Factor VIII mutants having increased half-life useful for treating  
PT hemophilia, comprise one or more amino acid substitutions in the A2  
PT and/or C2 domain of factor VIII -  
XX  
PS Claim 9; Fig 2A-B; 121pp; English.  
XX  
CC The invention relates to human factor VIII mutants comprising an amino  
CC acid substitution at one or more positions in the A2 domain and/or an  
CC amino acid substitution at one or more positions in the C2 domain.  
CC The invention also encompasses a factor VIII mutant which lacks a B  
CC domain (AAB8842). The factor VIII mutants have an increased half-life  
CC in the bloodstream. The A2 domain mutants exhibit reduced LRP-dependent  
CC (receptor-dependent) clearance of factor VIII, while C2 domain mutants  
CC have reduced receptor-independent clearance. The invention also relates  
CC to a method of using RAP (receptor associated protein), a protein which  
CC inhibits LRP (low density lipoprotein related protein), mediated ligand  
CC internalisation, to increase the half-life of factor VIII. The mutant  
CC factor VIII proteins, and nucleotides encoding them, are useful  
CC for treating haemophilia. RAP, LRP-binding RAP mutants or fragments, and  
CC nucleic acids encoding them may also be used in the treatment of  
CC haemophilia. In combination with a mutant factor VIII protein or DNA of  
CC the invention. The invention provides means of increasing the half-life  
CC of factor VIII by reducing its clearance from plasma. The present  
CC sequence represents a mutant mature human factor VIII which lacks a B  
CC domain.  
XX  
SQ Sequence 1424 AA:  
  
Query Match 96.4%; Score 7585; DB 22; Length 1424;  
Best Local Similarity 98.0%; Pred. No. 0;  
Matches 1423; Conservative 0; Mismatches 1; Indels 28; Gaps 1;  
  
QY 20 ATRRYVIGAVELSDVYQSDLGELPVDARPPRPKSPFRTSVYVKKTLFVEFTVHLEN 79  
DB 1 atrryvigavelsdvyqsdldgelpvdarpprpksfpfrrtsvvykkltfveltdhfn 60  
  
QY 80 IAKRPMPMGLIGPTIOAEVYDVTIVTLKNASHPVSLHAGVSYKMASGSAEYDDQTSQ 139  
DB 1 iakrppmgligptioaevydvitvltknashpvsilhagvsykmassgaeyddqtsq 120  
  
QY 140 REKEDKVPFGSGHTVYVOVLKENGPMASDPLCLTYSYLSHVDLVKDLNSGLIALLVCR 199  
DB 121 rekedkvfpfgshtlyvovlkenqpmasdpcltlysylishvdlvkdlnsgliallvcr 180  
  
QY 200 EGSIAKKEKQTQLKFIILPAVFDGSKSMHSTKNSLMODRBAASARAPMKHTYNGVYNR 259  
DB 181 egslakkekqtqlkfiiilpavfdgsksmhstknslmodrbaasarapmkhtvnygvnr 240  
  
QY 260 SLPLGICHRKSYVMHYIGMGTTPREVSHIFEGHTPLVRNHRQASLEISPTPLTAOTIL 319  
DB 241 slplgichrksyvmhyigmgttpevshifeghtplvrnhrqasleisptpltaotil 300  
  
QY 320 MDLQGFLLFCHISSHQHDMGMEAYVVDSCPEEPQLRMKNNEADYDDDLTSEMDEVYRF 379  
DB 1 mdlqgfllfchisshqhdmgmeayvvdscpeepqlrmknnedadydddltsemddevyrf 360

DB 301 mdlqgllfchissqhgdmeayvkvdscpeepqlrmkneeadydddltsemdvvrf 360  
QY 380 DDDNSPFIQIRSVAKKPKTWVHYIAAEEDDMYAPLYIAPDDRSVKSOYLNNGPQRIQ 439  
DB 361 dddnspfiqirsvakkpkptwvhyiaaeedwyparvlylapddrsvksqylngpqriq 420  
QY 440 RYKRVRFMAYTDETFKREAIQHESGILGPLGVEDTLTIFKQNASPPYNIYPIGI 499  
DB 421 rkykvrfrmaytdetfkreriaiohesgilgplgvedtltifkqnasppyniyipgi 480  
QY 500 TDVRPLYSRRLPKGVKHLKDEPILPGETFKYKWTVVEDGPTKSDPRLARYSSPVNME 559  
DB 481 tdvrplysrilpkgvkhlkdepilpgetfkykwtvvedgptksdprrlaryssfvnme 540  
QY 560 RDLASGLIGPLILICYKESVDORGNOIMSDKRNVLFSFEDENRSMYLTENIQRPLNPAG 619  
DB 541 rdlasgligplilicykesvdorgnolmsdkrnvlfsfedenrsmylteniqrlnpag 600  
QY 620 VOLBDEPFOASNIHNSINGYVFDLSQVCLHEVAWYILSIGAOTPLSYFSGYTPFKH 679  
DB 601 vlbedpfoasnihnsingyvfdlsqsvclhevaawyilsigaotplsyfsgytpfk 660  
QY 680 KMVEDTLTLTFPFGSETVFMSENPGLMILGCHNSDFERNQMTALIKVSSCDKMTGDYE 739  
DB 661 kmvedtltltfpgsetvfmsempglwilygchnsdfnrgmtallkvsckdntgdye 720  
QY 740 DSYEDISYLLSKNNALIEPRFSQNSRNPSTRQOFNATPVYLKRHREIIRTLQSQOE 799  
DB 721 dsyedisayllsknnaieprfsqnsrnpstrqofnatpvylkrhreirrtlqsqoe 752  
QY 800 EIDDDPDTISVMKEDDIDYEDENQSPRSFOKTRHFIYAVERLMDYSGSSPHVLRN 859  
DB 753 eidddpdtisvmkeddidyedenqsprsfoktrhfiyaavermldysgssphvlyrn 812  
QY 860 RAQSGSVQPFKKVVFQETDGSFTQPLYRGELNHLGLGPLYIAEVEDNIMVTFRMOAS 919  
DB 813 raqsgsvpfkkvfvfetdgsftqplyrgelnhlglyiaavednimvtfrmoas 872  
QY 920 RPYFYSLSIYEDDOGAGAPRKNFYKRNFTKTYFKVQNHMAPTDEPFCKMAATFSD 979  
DB 873 rpyfyslsiyeedogagaprknyfkrnftktyfkvnhmaptdpepfckmaatfsd 932  
QY 980 VDLKEDVHSGILGPLVLCNTTLPANAGROYTQOEALFPIPETKSWTYTENNERCR 1039  
DB 933 vldkedvhsgilgplvlcnettlanagroytqoefalfpietkswtytennercr 992  
QY 1040 APCNIQMEDPTFKENYRFHAIINGYIMDTLPGLVAADQRIKRWYLLSGNSNENIHSIFSG 1099  
DB 993 apcnlqmedptfkenyrfhaiingyimdtlpglvaadqrikrwyllsgnsnenihisifsg 1052  
QY 1100 HFTVVRKKKEEYKALNLYPGVEFTEVEMLPKSKAGIMRYECLIGHILNAGNSTFLVYSNK 1159  
DB 1053 hftvvrkkkeeykallnlypgveftevemlpkaskagimryeclighilnagnstflvysnk 1112  
QY 1160 CQPLGMAHGHIRFOFTASGOYQOMAPKLARLHYSGSINAMSTKEPFSWIKVOLLPMI 1219  
DB 1113 cqplgmasghirfdftasgoyqomapklarlhyssginamstkepfswikvolidpmi 1172  
QY 1220 IHGIRTOGARQKFSLSYISOFIIMYSLDGKMWQTYRGNSTGTLVFFGANDSSGIRKNIF 1279  
DB 1173 ihgirtogarkfsslyisofiiimysldgkwmqtyrgnsgtltlvffgandssgirknif 1232  
QY 1280 NPPIIARYIRLHPHYSTRSLRMLWGLDNLNSGSMPLGMSKALISQAQITASSYFTNMF 1339  
DB 1233 nppliiaryirhphystslrmlwglndlnsgsmplgmsskalisqaqitassyftnmf 1292  
QY 1340 ATWSPSKARLHLIOGRSNAMRPQVNNPKEMLOVDQKTMKVTGYTQGVKSLTSMYKEF 1399  
DB 1293 atwspskarhlhiogrsnamrpqvnnpkewlvdqtkcmkvtgytgvkssltsmykef 1352  
QY 1400 LISSSDGCHOWTLTFQNGKVKVFGQNDSTFPVNVSLDPLRLRYLRINPQSWWQIALR 1459  
DB 1353 lissdgchowltltfngkvkvfgngdsfprvvnslpdlrlrylrinpqswwhqialr 1412







Qy	1340	ATWSPKALHLHOGKSNANRPVNNREKWTQVDFQRTMAKYGTGTTGCKSLTSMYKEF	1399
Db	1294	atwspkarlhlhgksnawrpvnnpkewlqvdfqktmkvtgvtlgvkslltsmykelf	1353
Qy	1400	LTSSSDDGHWTFEPONGKVKVFGQNDSPFPVNSLDPPLRLRYLRHFOQSWHOTALR	1459
Db	1354	ltsssdgnhwtfepngkvkvfgqndsfcpvnsldpplrlcrylrhpgswvhtalr	1413
Qy	1460	MEVLGCEADOLY 1471	
Db	1414	mevlgceaqdly 1425	
RESULT 11			
AAp80268	ID	AAp80268 standard; protein; 1424 AA.	
XX	AC	AAp80268;	
XX	DT	10-OCT-1990 (first entry)	
XX	DE	Modified factor VIII:C sequence with the R740-D1658 deletion.	
XX	KW	Modified factor VIII:C; maturation polypeptide; haemophilia;	
XX	OS	blood coagulation; RD deletion; procoagulant.	
XX	PN	Homo sapiens.	
XX	PD	WO8800831-A.	
XX	PF	11-FEB-1988.	
XX	PR	31-JUL-1987; 87WO-US01814.	
XX	PA	01-AUG-1986; 86US-0893375.	
XX	PI	(BIOJ ) BIOGEN NV (PASE/).	
XX	DR	Pasek MP;	
XX	DR	WPI; 1988-049866/07.	
XX	PS	N-PSDB; AAN80447.	
XX	PT	New DNA sequences encoding modified factor VIII:C - with deletion of DNA	
XX	PT	encoding maturation polypeptide, useful for high yield transformation.	
XX	PS	Claim 3; Page 60-61-62-63; 97PP; English.	
XX	CC	The RD deletion removes the DNA from Ser 741 to Ser 1657.	
XX	CC	A major part of the sequence encoding the maturation polypeptide of	
XX	CC	factor VIII:C is deleted, i.e. Gln 744 - Asp 1563.	
XX	CC	The full length factor VIII:C cDNA has two changes with respect to the	
XX	CC	published sequence (EPO application 160457):	
XX	CC	CTG to CTA at Leu 242 and TTC to CTC change at amino acid residue 1880	
XX	CC	(Phe to Leu). The product is produced in approx. 20 times higher	
XX	CC	yields than previous recombinant produced factor VIII:C and are more	
XX	CC	easily purified. The peptide is used for treating haemophilia A, both	
XX	CC	acute and prolonged bleeding.	
XX	CC	See also AAN80444 and AAN80446.	
XX	CC	Sequence 1424 AA;	
XX	CC	Sequence 1424 AA;	
Qy	Query Match	96.2%; Score 7571; DB 9; Length 1424;	
Db	Best Local Similarity	97.8%; Pred. No. 0;	
Matches	1420; Conservative	1; Mismatches 3; Indels 28; Gaps 1	
Qy	20	ATRRVYLGAVELSMQYMOGDLGELPVDARFPPRPVRSKSPFNTSVYKKTLFVEFTVHLNF	79
Db	1	atrrylgavelswdymgsdldgelpydarfpprvksfntsvykkcllfeftchln	60
Qy	80	IAKPRPFWMGLLGPTIQAEVYDVTVVITLKNMASHPVSLHAGVSVYKASEGAEYDQTSQ	139

Dd	61	l	a	k	r	p	p	m	g	l	i	p	r	i	q	a	e	y	d	v	w	i	c	k	m	a	s	i	p	s	l	a	v	a	y	s	y	k	a	s	e	a	e	y	d	d	t	s	q	120						
Qy	140	r	e	k	e	d	k	e	f	e	c	g	s	h	t	y	w	o	v	i	k	e	n	g	p	a	s	p	l	c	t	i	v	s	y	s	h	v	d	w	i	d	n	s	g	i	a	l	w	c	r	199				
Dd	121	r	e	k	e	d	k	e	f	i	p	s	s	h	t	y	w	q	v	i	k	e	n	g	m	a	s	p	c	i	t	y	s	i	s	h	v	d	i	v	d	i	n	s	g	i	a	l	w	c	r	180				
Qy	200	b	e	s	l	a	k	e	r	t	o	t	l	i	k	e	f	i	l	e	f	a	v	e	d	e	k	s	w	s	e	t	k	n	s	i	m	o	r	d	a	s	a	r	a	m	p	k	m	t	v	n	c	r	259	
Dd	181	e	g	s	i	a	k	e	t	h	e	i	k	f	i	l	f	a	v	i	d	e	g	k	s	w	h	e	t	k	n	s	i	m	q	d	a	s	a	r	a	w	k	m	t	v	n	c	r	240						
Qy	260	s	i	p	g	i	c	h	r	k	s	y	w	h	y	i	a	m	g	i	t	r	e	v	n	s	i	f	e	g	h	t	i	v	r	h	r	a	s	e	i	s	p	i	r	t	a	o	t	i	l	319				
Dd	241	s	i	p	g	i	c	h	r	k	s	y	w	h	y	i	a	m	g	i	t	r	e	v	n	s	i	f	e	g	h	t	i	v	r	h	r	a	s	e	i	s	p	i	r	t	a	o	t	i	l	300				
Qy	320	m	d	i	c	e	f	f	e	c	i	s	s	h	o	d	g	m	a	y	k	v	a	y	s	c	e	e	p	o	l	r	k	m	k	n	e	e	a	d	e	x	d	d	i	t	s	e	m	d	y	a	f	379		
Dd	301	m	d	i	c	e	f	f	e	c	i	s	h	o	d	g	m	e	a	y	k	v	a	y	s	c	e	e	p	q	i	m	k	n	e	e	a	d	e	x	d	d	i	t	s	e	m	d	y	a	f	360				
Qy	380	d	d	d	n	s	p	f	i	o	i	r	s	v	a	k	k	p	k	t	w	h	y	i	a	e	e	d	w	i	a	r	y	a	r	p	d	r	s	k	y	o	q	i	n	g	p	o	r	g	439					
Dd	361	d	d	d	n	s	p	f	i	q	r	s	v	a	k	k	p	k	t	w	h	y	i	a	e	e	d	w	i	a	r	p	d	d	r	s	k	y	q	i	n	g	p	o	r	g	i	g	420							
Qy	440	r	k	y	k	v	r	m	a	t	t	d	e	f	e	r	t	r	e	a	i	o	h	e	s	i	c	e	r	p	l	l	x	e	y	e	g	t	l	i	f	r	n	o	s	r	p	y	n	i	p	h	c	i	499	
Dd	421	r	k	y	k	v	r	m	a	t	e	f	i	k	r	e	a	i	q	h	e	s	i	g	p	l	i	y	e	g	v	d	i	l	i	f	k	n	a	s	r	p	y	n	i	p	h	c	i	480						
Qy	500	t	d	v	r	p	l	y	r	r	l	r	p	k	v	h	i	k	p	e	r	p	i	l	p	g	e	i	f	e	y	k	m	t	v	i	e	o	p	k	r	s	p	r	c	t	r	r	y	s	s	p	v	n	e	559
Dd	481	t	d	v	r	p	l	y	r	r	l	r	p	k	v	h	i	k	p	e	r	p																																		



Db 1113 cqlplmasghlrdqfqlasggvgqwapklarlhyssinawskpefswikvdlapml 1172  
Qy 1220 IHGIRTOGAROKFSSLYISOFITMYSLDGRKWOTYRGNSTGTLMVFEFVNDSSGIKHNIF 1279  
Db 1173 ihglttgargqfsslylsqflimysldgkwtqlyrgnstgltlmvffgnvdssgikhnif 1232  
Qy 1280 NPPIIARYIRLPHTHYSIRSTRIMELMGCDLNSCSNPLMESKASIDAOITASSYFTNMF 1339  
Db 1233 npplliarylrhlphthysirstrclmelmgcdlnscsmpmgmeskasidagqlassyftnmf 1292  
Qy 1340 ATWSPSKARLHOGRSNMRPOVNNPKELQYDFOKTMKVTGTTQGVKSLTSMYKKEF 1399  
Db 1293 atwspskarlhlgqrsnawrvpnmpekewlgydfqktmvtgvtlqgvkslltsmyvkef 1352  
Qy 1400 LISSQDGHOWTLFPQNGKVKVFOGNSDFTPVNSLDPPLLTRYLRIHPOSVWQIALR 1459  
Db 1353 lissqdgghqwtlffqngkvkvfgngqdsftpvnsldpplltrylrihpgsvwqhialr 1412  
Qy 1460 MEVLGCEADLY 1471  
Db 1413 mevlgceadly 1424  
RESULT 12  
AAP91169  
ID AAP91169 standard; protein: 1424 AA.  
XX  
AC AAP91169;  
XX  
DT 26-JUN-1990 (first entry)  
XX  
DE Sequence of 740 Arg-1649 Glu human Factor VIII:C.  
XX  
KM Human Factor VIII:C; Ad.RE.neo: 740 Arg-1649 Glu Factor VIII:C;  
KW haemophilila A.  
XX  
OS Homo sapiens.  
XX  
PN EP306968-A.  
XX  
PD 15-MAR-1989.  
XX  
PE 09-SEP-1988; 88EP-0114769.  
XX  
PR 08-APR-1988; 88JP-0085454.  
XX  
PA (KAGA ) CHEMO-SERO-THERAP (TELU).  
XX  
PI Sugiyama T, Masuda K, Tajima Y, Yonemura H;  
XX  
DR WP1: 1989-078467/11.  
XX  
DR N-PSDB; AAN90654.  
XX  
PT Prodn. of recombinant human Factor-VIII-C -  
PT using animal cells transformed with a vector contg. the gene for  
PT Factor VIII:C and a promoter  
XX  
PS Fig 1(1) - 1(13); : 32pp; English.  
XX  
CC Arg-740 of the carboxyl terminus of the H chain is  
CC directly bonded by a peptide bond to Glu-1649 of the amino terminus of  
CC L chain. A prefd. expression vector used to transform animal cell so  
CC that they produce human Factor VIII:Cis plasmid Ad.RE.neo.  
CC The expression vector has at least one promoter upstream of AAN90654.  
CC The transformants can constantly and continuously produce human Factor  
CC VIII:C in high yield on a commercial scale. The human Factor VIII:C so  
CC produced is considered to corresp. to the smallest species of active and  
CC intact Factor VIII:C molecules in the human blood plasma. It is useful  
CC for treating haemophilila A patients.  
XX  
SO Sequence 1424 AA;

Query Match 96.2%; Score 7567; DB 10; Length 1424;  
Best Local Similarity 97.8%; Pred. No. 0;  
Matches 1420; Conservative 0; Mismatches 4; Indels 28; Gaps 1;  
Qy 20 ATRRYTLGAVELSWDMQSDLELYDARFPPRVKSPSPFNSTVYKKTLYVEFTVHLFN 79  
Db 1 acrrylgavelswdmqsdlqelpydarfprrvksfpfnstvykktlyvetfdhfn 60  
Qy 80 IAKRPPMMGLIGPTIOAVYDTVTYITLKNMASHHVSJHAWGVSWKSEGAEDDQTSQ 139  
Db 61 iakrppmmglligptlgaevydtvtltnkmaslhvsjhaavgvswkasegaedqdtsg 120  
Qy 140 REKEDDKVEFGSSHTYVWQVLKENGPMASDPLCLTYSLSHVDLYKDLNSGLIGALLVCR 199  
Db 121 rekeddkvrfpgshcywqvlkengpmasdpcltysylshvdlvkdlngllgallvcr 180  
Qy 200 EGSIAKKEKQTLHKFTLLFAVFEDEKSMHSEKNSLMDRDAASRAAPKMHYNGVYNR 259  
Db 181 egslakkekqtlhkftllfavfdegkswnselknsldrdaasarapkmhlyngvynr 240  
Qy 260 SLPLGLICHRKSVYWHVIGMGTTPVHSTFLBGAHFLVRNHRQASLEISPTFLTAOTIL 319  
Db 241 slpllgicchrksvywhvigmgttpevhsflbghflvrnhrgasleisptfltaqtl 300  
Qy 320 MDLGOFLFCHTSSHQHDMEAIVKVDSCPEEPQLRMKNNEAEDYDDLTDEMDVYRF 379  
Db 301 mdlgqflfchthshqhdmeayvkvdsceepqlrmkneeaedyddltdeemdvyrf 360  
Qy 380 DDNSPSFTQIRSVAKKHKHTVWHYIAAEEDMDVAPLYLAPDDSYKSOYLNNCPQRIQ 439  
Db 361 ddnspsftqirsvakkhkhtvwhylaeeedmdvaplylaphdtsyksqylmnpqrlq 420  
Qy 440 RYKRVKFMAYTDEFTFKTEAIOHESGLIGPLLYGEVDTLLIIFKNASPPVNIYHGI 499  
Db 421 rkykvrkfmaytdeftkteaiohesgllgpllygevdtlllflknasppnylhybgi 480  
Qy 500 TVVRPLYSRRLPKGVKHLKDEPILPGEIFKRYKWTVJEDGPTKSDPCLTRYSSFYVME 559  
Db 481 tvvrplysrllpkgvkhlkdpilpgeifkykwetvtedgptksdpccltryssfyme 540  
Qy 560 RDLASGLIGPLLICKESVDQNGNOIMSDKRVILFSVDENRSXYLTENQRLPAPAG 619  
Db 541 rdlasgllgpllickesvdqngnomsdkrvilfsvdenrsxylytenqrlplpapg 600  
Qy 620 VOLEDPEFGASNIMHSINGYVDSLOLSVCLHEVAYWYILSIGAQTDFLSFFSGFYFKH 679  
Db 601 vledpefgasnimsingyvfdsqsvclheavaywylsigqtdtlfssfygfkfh 660  
Qy 680 KMYVEDTTLTLPFSGEIVFMSMKNPGLMILGCHNSDFRNKMTALLKVSCKDKNTGDYE 739  
Db 661 kmvyedttltlpfsgetvfmsmnpglwllgchnsdfnrgmtallkvsckdntgdye 720  
Qy 740 DSEYDISAVILSKNNAIPRSPSSQNSRHPSTRQKFNATPVYLKRHOEITRTLOSQOE 799  
Db 721 dseydisavilsknnaip-----reltrtlqsdqe 752  
Qy 800 EIDYDQTSIVEAKKEDFDIYDEENQSPRSFOKTRHYFAVERLMDYGMSSSPHYLRN 859  
Db 753 eidyddtsiveakkedfdiydedenqsprsfqktrhyfliaaverlwdygmsssphyln 812  
Qy 860 RAQSSVPOFKKVVPOEFTDGSFTOPLYRGEINELHGLGPLYIRAEVEDNIMVTRNOAS 919  
Db 813 raqssvpofkkvvpoeftdgstftoplyrgeinelhglgplyiraevednimvtrfnqas 872  
Qy 920 RYRSFYSSLISYEEDQROGAERKNFVFNPKTYFKMKVQHMMAPTKEPFCKAMAYFSD 979  
Db 873 rrsfysslysyeeqrgaeprrknfvkpncktyfwkvqhmmpckdeickakawayfsd 932  
Qy 980 VOLEKDVHSGILGIPLLVCHTNTLPAHAGROYTVOEFAFLFTFDETKSMYTTENMERCR 1039  
Db 933 vlekdvhsqilgpllvchtntlnpahagroytvogefalfftfdeetskmyttemerncr 992  
Qy 1040 APCNTIOMEDPTEKENYRPHAINGYIMDTLPGLVMAQDQIRIMWYLLSMGSENHISHIFSG 1099











FT /note="Signal peptide"  
FT Protein 20..2351  
FT /note="Mature factor VIII:C"  
FT Misc-difference 671  
FT /label= phe652Hls  
XX  
XX MO9703191-A1.  
XX  
XX 30-JAN-1997.  
XX  
XX 28-JUN-1996; 96WO-0511013.  
XX  
XX 11-JUL-1995; 95US-0001030.  
XX  
XX (CHIR ) CHIRON CORP.  
XX  
XX Hung DT;  
XX  
XX WPI; 1997-119047/11.  
XX  
XX Factor VIII:C analogue - modified to comprise Factor V A or C domain  
XX or subdomain, for increased stability or activity  
XX  
XX  
XX Claim 6; Page -: 45pp; English.  
XX  
XX The sequences given in W10590-92 represent active Factor VIII:C (F8C)  
XX C polypeptide analogues. These analogues comprises a native F8C  
XX C polypeptide modified to comprise the presence of a Factor V A and/or  
XX C domain and/or subdomain. The F8C polypeptide analogues, alone or in  
XX combination, can be used for the prevention or treatment of an active  
XX F8C deficiency, i.e. haemophilia A and other blood clotting disorders.  
XX The analogue can also be used as an immunogen for antibody production.  
XX The analogues can have an increased plasma half-life, or specific  
XX activity. This sequence is not given in the specification and is  
XX based on the sequence derived from Genbank ref. K01740.  
XX  
XX  
XX Sequence 2351 AA:  
SQ  
Query Match 94.28; Score 7416; DB 18; Length 2351;  
Best Local Similarity 62.5%; Pred. No. 0;  
Matches 1470; Conservative 1; Mismatches 0; Indels 880; Gaps 1;  
QY 1 MOELSTCFPLCLRFCEFSATRRYYLGAVELSMQSDGELPVDARPPRPVPSPPFN 60  
DB 1 mqlstcfplcllrfcefsatrryylgavelswymgsdlgelpvdarfprrvpksfpfn 60  
QY 61 TSVYKKTLEVEFTVHLFNIAKPRPPMGLGPTIOAEVYDTVVITLKNNASHPVSLHAV 120  
DB 61 tsvykktlfevftvhlfnlajakprppmglgptioaevydtvvitlknashpvsihav 120  
QY 121 GVSYWKASGEAEYDDQTSQREKEDKVPFGSGSHYVNOVLKKNPMAADPLCLTYSLSH 180  
DB 121 gvsywkasegaeayddqtsqrekedkvfpqsgshyvnovlkennpmasdplcltysylsh 180  
QY 181 VDLAKDINSGLIGALVYCRGSLAKEKQTLHKFILLFAVDEKSWHSEFKNSIMODRD 240  
DB 181 vdlakdinsgligallvycrgslakektqlhkfillfavdeqskshsetkslmqdrd 240  
QY 241 AASARAPKMHYNGVYVNSLPGILGCHRSVYWHVIGMTPEVHSIFLEGHTFLVRNH 300  
DB 241 aasarapkmhlyngvynsrslpgilgchrsvywhvigmtpetvhsifleghftlvrn 300  
QY 301 RQASLESPITFLAOTLLMDLGOFLFCHTSSHQHDMGEAYVVDSCPEEPQIRKMNNE 360  
DB 301 rqaesletpitflaotllmdlgofllfchthsshqhdmgeayvvdscpeepqirkmne 360  
QY 361 EAEYDDDLTDESEMDVYRFDDDNSPSFIQIRSVAKKHPKTMVHIIAEEDMDVAPLVLA 420  
DB 361 eaeydddltdesemdvyrfdddnspsfioirsvakkhpktmvhiiaeedmdvaplvla 420  
QY 421 PDDRSYKSYLNNGPORIGRKYKVRPMAYTDETFKTREAIQHSGLIGPLLYGEVGDTL 480  
DB 421 pddrsyksylnngporigrkykvrpmaytdefktrreaiohsiglpllygevgdtl 480

DB 421 pddrsyksylnngpqrfgrtkykvrfmayldetfktreaiqhesgllgpllygevgdtl 480  
QY 481 LIIFKNQASRRYNYIRPGITMDVRRLYSRRLPKGYKHLKDPFILLGELTFKRYWYTFVEGP 540  
DB 481 liifknqasrryinyirpgitmdvrRLYsrRLpkyKHLkDPfIlLgELtFKRYWYtFVEGP 540  
QY 541 TRSDPCLTRYSSSFVNMNERDLASGLIGPLICYKESVDQGNQIMSDRRVILFSVDE 600  
DB 541 trsdpcltrysssfvnmnerdlasgligplicykesvdqgnqimsdrrvilsfvde 600  
QY 601 NRSWYLTENIQRLPNPAGVQLEDEFEQASNIMHSINGVYDLSQLSVCLHEVAYWYILS 660  
DB 601 nrswyLTENIQRLPNpAGVQleDEFEQASnImHsIngVYdLSqLSVclHeVAYWYILs 660  
QY 661 IGAOTDFLSVPSGTYTRKHKKVYEDTLTFPFSGETYPMSENENGLWLGCHNSDFPRKG 720  
DB 661 igaoTDFLSVpSGTYtrKHKKVYEDTLTFpFSGETYPMsENENGLWlGCHNSDFPRKG 720  
QY 721 MTALLKVSQCDKNTGDYEDSEYEDISAVLSKNNAIERPSFSONSRHPSTRQKOFNAT-- 778  
DB 721 mtaLLKVSQcdKntGDYedSEYEDISaVLSKnnAIERpsfSONsrHPSTRqKOfNat-- 778  
QY 779 ----- 778  
DB 781 pendiektpwfahrtmpkignvssdlmlrlqsptphglsldiqeakyetfsddps 840  
QY 779 ----- 778  
DB 841 pgaidsnslsemhffpqjlhsgdmvftpesqqlrlneklgtatelaKklqkvsst 900  
QY 779 ----- 778  
DB 901 smnlstcphnlaagrdnteslqpsmpvnydsqldtllfgkxspstsesgplslsee 960  
QY 779 ----- 778  
DB 961 ndskllsajmsqesswgnvssesagrllfkqkrahgpalltkdnalfkvsislktn 1020  
QY 779 ----- 778  
DB 1021 ktsnsatnrkthldpslllenspsvqnllsdeffkvtpllhdmimdnatalrl 1080  
QY 779 ----- 778  
DB 1081 nhmsnkttskmemvqkkesgplpdaqnpdmsffkmlflpesarwlrthgknsinsg 1140  
QY 779 ----- 778  
DB 1141 qgpepkqlvsigpek svegqnflsekknvvvgkgeftkdvglkemvfpssrnlfltnldn 1200  
QY 779 ----- 778  
DB 1201 lhennthngekkigeleketlligenuvlpqjlhtvcgknfmknllflstrgnvegsye 1260  
QY 779 ----- 778  
DB 1261 gayavpvgdfrrslnsncnrkktahabfskgeeenlsglgnqtkqivekyacttrispnt 1320  
QY 779 ----- 778  
DB 1321 sqgnfvgrskralkgfrrlpleetelekrllvddtscqsknmkhlprstltqdyneke 1380  
QY 779 ----- 778  
DB 1381 kgaletgspldccltrshslpqaansplpjakvsfpirplylrrvlfgdnsshlpaasy 1440  
QY 779 ----- 778  
DB 1441 rkksdyqgeshflgskakmnlslaitltemtdqrevgsldgtsatnsvlykkventvlp 1500  
QY 779 ----- 778  
DB 1501 kpdlptkskvellpkvhiykdalfpctetsngspghidlvsgsligqtegaikwneanrp 1560



QY 779 ----- 778  
Db 1561 gkvpflvatesasaktpskllldplawdhhygtqrlpkeewksqekapektafkktllsl 1620  
QY 779 ----- PVLAKRIQRETRTTLQSDQPE 800  
Db 1621 nacesnhalaieinegnqkpelevtwakqgarterlcsqnpvllkrlqreltrtltsdsqee 1680  
QY 801 IDYDITTSVEAMKKEBFDIYDEDEENQSPRSFOKTRHRYFLAAVERLMDYGMSSPHYLRNR 860  
Db 1681 ldyddtlsvemkkedfdldydedengsprsfckltrhyflaaverlwdgmssphvlmr 1740  
QY 861 AQSQSVPOFKVVFQEFDTGDSFTQPLRYGELNEHLGLGAPYRAVEDNINWTFRNQASR 920  
Db 1741 agsgsvpqkfvfgeftdgsftqplrygelnehlglgpyraevdedlmvtfingqae 1800  
QY 921 PYSFYSSLSISEEDOROCABERKNKFNKFNKTYFMKVOHNAHPKDEDCAMAYFSVY 980  
Db 1801 pysfyslslsyeedqrgaepkntvkpnetklytwkvqhmapktdedckawayfsdv 1860  
QY 981 DLEKDVHSGILGPLVCHTNTLNPAGHROVYVOEFALFETIDETKSMYFTENMRNCRA 1040  
Db 1861 dlekdvhsgllgpllvchntlnpnhgqvqvgefalffltideltksyftemnerncra 1920  
QY 1041 PCNIQMEDPTFEKENVRFHAINGYIMDTLPGLVMAODRIRWYLSMGSNENHSHFSGH 1100  
Db 1921 pcniqmedptfkenyrfhaingyimdtlpglvmaqdrfrrwyllsmgsnenhshfsgh 1980  
QY 1101 VFYTRKKEEYKMALYNLYPGVFETVEMLPKSKAGIWRVECLIGEHHAAGSTFLVYSKNC 1160  
Db 1981 vfytrkkeeykmalynlypgvfetvemlpksglwrvecllgehlhagmstflvysnkc 2040  
QY 1161 QTPLGMSAGHARDPOITASGQYGQWAPKLARLHYSGINAMWSTKPPFWIKVDLAPMT 1220  
Db 2041 qtplgmssghlrdpogitassgqygqwapklarlhysgsinawstkepfwixkvdllapmt 2100  
QY 1221 HGKIQGAKRQKFFSLYISQFTIMSLDGKMWOTYRGNSTGTLMEFGVNDSSGKIHNFN 1280  
Db 2101 hgkltgarkqkfsslylsqftilmsldgkwmotyrgnstgtlmvffgnvdsqikhnf 2160  
QY 1281 PPIIARIYRLHPHTHYSRSTRLMELMGCDLNSCMLMESKASIDAOQTASSTYTNMRA 1340  
Db 2161 ppiariyrlhphthysrstrlmelmgcdlnscmplmeskasisdaqtlcassyltmra 2220  
QY 1341 TWSPSKARLHLOGRSNAMRPQVNNPKEWLQVDFOKTMKVTGTOGVKSLTSMYKFEFL 1400  
Db 2221 twpskarlhlqgrsnawrpqvnnpkewlqvdfoktmkvgttqgvkslltsmvykefl 2280  
QY 1401 ISSSQDGHQWTLFFQNGKVKVFGQNDSTFPVNSLDPPLLTRYLRIHPQSVNHQIALRM 1460  
Db 2281 isssqdghqwtlffqngkvkvfgqndstfcpvnsldpplltrylrlhpqsvnhqialrm 2340  
QY 1461 EVIGCEAODLV 1471  
Db 2341 evlgceaqdly 2351

RESULT 15  
ID AAM13496 standard; Protein: 2351 AA.  
AA13496:  
AC AAM13496:  
XX  
XX  
DT 03-DEC-1997 (first entry)  
DE  
XX Factor VIII:C (Arg1689Lys).  
XX Factor VIII:C (Arg1689Lys).  
XX Factor VIII:C; analogue; substitution; Factor VIII:C deficiency;  
XX Factor VIII:C; analogue; substitution; Factor VIII:C deficiency;  
XX Factor VIII:C; analogue; substitution; Factor VIII:C deficiency;  
OS Homo sapiens.  
XX  
XX

Key Location/Qualifiers  
FT Misc-difference 1708  
ET /Label= Arg1689Lys  
PN W09703194-A1.  
XX  
PD 30-JAN-1997.  
XX  
PF 09-JUL-1996; 96WO-US11441.  
XX  
PR 11-JUL-1995; 95US-0001029.  
XX  
PA (CHIR ) CHIRON CORP.  
PI Burke RL, Rosenberg S;  
XX  
XX WPI: 1997-119049/11.  
DR N-PSDB; AAT61548.  
XX  
PT Factor VIII:C analogue with Arg 1689 substituted by Lys - useful for  
PT preventing and treating a Factor VIII:C deficiency, particularly  
PT haemophilia A  
XX  
PS Claim 1; Page -: 46pp; English.  
XX  
CC This sequence represents an active Factor VIII:C polypeptide analogue.  
CC The analogue comprises a native Factor VIII:C polypeptide that is  
CC modified by substitution of the Arg residue at position 1689 with  
CC Lys. The polypeptide analogue can be used to prevent and treat a  
CC Factor VIII:C deficiency in a mammal, esp. haemophilia A. This  
CC sequence is not given in the specification and is based on the  
CC sequence given in Genbank Ref. K01740.  
XX  
SQ Sequence 2351 AA:  
Query Match 94.2%; Score 7416; DB 18; Length 2351;  
Best Local Similarity 62.5%; Pred. No. 0;  
Matches 1470; Conservative 1; Mismatches 0; Indels 880; Gaps 1;  
QY 1 MOELSTCFEFLCLRFCSATRRRYLGAVELSMWDMOSDLGELPVDARFPFRVPSFPFN 60  
Db 1 mqlstcfefllclrfcsatrrrylgavelswdmgsdglpvdarfprvpksfprfn 60  
QY 61 TSVVYKKTLEVEFTVHLFNIAKPRPMGLGPTIOAEVYDTVYITLKNMASHPSLHAY 120  
Db 61 tsvvykktlveftvhlfniaakprpmmglgptioaevydtvyitlknmaspshpslhav 120  
QY 121 GVSYKASBGAETVDTQTSQREKEDKVPFGSGSHYTWQVYLKENGPMA SDPLCLTYSLSH 180  
Db 121 gvsywkasegaeayddqtsqrekedkvfpgsgshytwqvlkengpmasdplcltyslsh 180  
QY 181 VDLVNDLNSGLIGALVYREGSLAKEKTOPLHKFTLPLFAVDEGSMHSEKNSLMDRD 240  
Db 181 vdlvndlnsglilgalvyrsgslakektpplhkftlplfavdegsmhseknslmrd 240  
QY 241 AASARAWPKMHTVNGYVNRSLPGLIGCHRSYVHWVIGMTTPEVHSIFLEGHTFLVRNH 300  
Db 241 aasara wpkmtvngyvnrslpgligchrsyvhwvigmtpetvhsifleghtflvrnh 300  
QY 301 ROASLEISPTFTFLTAQTLLMDLGFLLFCHISSHOHGMEAYVYVDSCEPPEQLRMKNE 360  
Db 301 rgasleisptftfltaqtlldmldgflfchisshohgmeayvrvdscpeepqlrmkne 360  
QY 361 EAEDYDDLTOSEMNVYRFPDDNSPFOISVAKKHKPKTWVHIIAAEEEDMDVAPLVLA 420  
Db 361 eaedyddldtosemnyvrfpddnsfpfoisvakkhkpktwvhiiiaaeeedmdvaplvla 420  
QY 421 PDDRYSKSOYLNNGPORIGRKYYKKVFMAYTDEFTKTEAIOHSGLIGPLLYGEVGDTL 480  
Db 421 pddr ysksoylnngporigrkyykkvfmaytdef tktteaiohs gligpllygevgdtl 480  
QY 481 LTIENQASRPYNIYPHGITDVRPLYSRRLPKGVKHLKDPFLPGLGELFKKMYTVTEDGP 540



|||||  
Db 481 lllfngsrrpniyphglttdvryrrrlpkyvnhlkdfprrllpgeikkyktvveqgp 540  
QY 541 TKSDBRCLTRYSSVFVNERDLASGLIGPLLCYKESVDORQNOIMSKRNVILFVSFDE 600  
Db 541 tksdprcltrrysffvmerdlaaglllqplllcykesvdqrgnqmsdkrvnllfsvde 600  
QY 601 NRSWLTENIRFLPNPAGVOLEDEPEQASNTMHSINGVPSILOSLCHVAAWYLLS 660  
Db 601 ||||| 660  
QY 661 IGAQDTLVSFESGYTFKHKWYEDTLTLFPFSGETVFMSEMENPGILWILGCHNSFRNRG 720  
Db 661 lgaqdtllsvffsgrtfkhwyeetllclfpsetvfmsemenpgllwlgchnsfrnrg 720  
QY 721 MTAALKVSCDKNKGDYEDSTEDISAYLLSKNNAIEPRFSQNSRHPSTROKOFNAT-- 778  
Db 721 mtaalkvsscdkntgdyedsedlsayllsknaieprsfqsgrhpsrtqkfnaatl 780  
QY 779 ----- 778  
Db 781 pendlektcdwfhartcpmpkignvssdqlmlllqspctphglsldlgeakyetfsdds 840  
QY 779 ----- 778  
Db 841 pgaldsmnslemthfrpqllhsgdmvftpesglqlrlnekigttaateklkldfkvest 900  
QY 779 ----- 778  
Db 901 snnllstlpsdnlaagtntsslgppsmvhydsqldtllfgkkespltesgplisee 960  
QY 779 ----- 778  
Db 961 ndsklllesglnmagesswgknvsslesgrllfkgrahgpalltkdna lfkvssllktn 1020  
QY 779 ----- 778  
Db 1021 ktsmsatnrkthidgplllienspsvwgnllesdtefkvtprllhdmlmdknaatalrl 1080  
QY 779 ----- 778  
Db 1081 nhmsnkttskrmemvgqkgegrlpdpdaqnpdmsffkmlflpesarwlrthgknslnsg 1140  
QY 779 ----- 778  
Db 1141 qgspkqlvsjgpeksvegnflseknkvvgkgeftkdvgllkemvfpssrnlfltnldn 1200  
QY 779 ----- 778  
Db 1201 lhemthngekkigeeketlligenvlpqllhtvtgktnfmknlflstrgnvesgys 1260  
QY 779 ----- 778  
Db 1261 gayapvlqdfslndstnrctkhtahfakgeeenleglgnqtkqivekyacttrispnt 1320  
QY 779 ----- 778  
Db 1321 sqgnfvtrskrralkqfllpleetelckriivdctstqwmknhltpstlltqldyneke 1380  
QY 779 ----- 778  
Db 1381 kgaltqpslsgcltrshsldpqnrsplakvssfslrplyltrlvlfqdnshpaasy 1440  
QY 779 ----- 778  
Db 1441 rkkdsyvgesshllqgakknlsajalllemtgddqrevgslyatsatnsvtkykventvlp 1500  
QY 779 ----- 778  
Db 1501 kpdlpkrtsqkveellpknvlyqkdllfprctsnsgpghldlvegsllqtegalakwneanrp 1560  
QY 779 ----- 778

Db 1561 gkvflrvateessaktprskllldplawdnhygrqdpkeekwsgekspektrafkxkddtllsl 1620  
QY 779 -----PVLKRRHOREITRPTTLOSDOBE 800  
Db 1621 nacesnhaalaenegnkpellevtwakggrterlcsqnpvplkrrqreltrlltqsdgee 1680  
QY 801 IDYDITISVEMKKEDFDIYDEDEENOSPRSFOKKTRHFFYAAVERIMDVGMSSPVLRNR 860  
Db 1681 ldydttisvemkkedfdlydedenpskpsfgkkltnhflaaaverlwdygmssspvllnr 1740  
QY 861 AQSGSVPQFKVVFQEDTDSFTQPLYRGELNEHLLGLAPYIRAEVEDNIMVTFRNOASR 920  
Db 1741 aqsgsvpqfkvvfgefdgstfqlpyrgelnehglllqpylraevdnimvtfngsar 1800  
QY 921 PYSFYSLSIYEDDROGAEPKRNKFNKNETKTYTWKVOHMAPTKDEFDCKAMAYFSDV 980  
Db 1801 pysfyslisyeeedqrgaepkfnkfvkpneltfkwkvghmapkdefdckawayfsdv 1860  
QY 981 DLEKDVHSGILIGPLLVCHTNTLNPAGROVTOEFALPFTIPDETRKSWTFENMERNCRA 1040  
Db 1861 dlekvhnsglllqpllvchntnlnpahngrqvvtvgefallftllfdektswyftenmerncra 1920  
QY 1041 PCNIOMEDPTFEKENYRFAINCYIMDTLPGLYMAODORIRWYLLSMGSNENIHSIHFSCH 1100  
Db 1921 pcniomgedptfkenyrfhaingyimdtlpglymadqgrlrvyllsmgsnenihsihfsg 1980  
QY 1101 VFTVRRKEEYKALYNLPGVEYVEMLPKAGIMRVECLISEHLLHAGNSTFLFLYYSNKC 1160  
Db 1981 vftvrkkeeykmalynlpgvfevtemlpjskagjvrvecllgehlhagnstlfllyysnkc 2040  
QY 1161 QTPGLMASGHIRDPOITRSGOVGOMAPKLARLHYSGSINAMSTKBEFSIKYDILLAPMI 1220  
Db 2041 qtpglmasghirtdqitltsagqy9gwapklarlhysgslnawstkepfswlxkdllapmli 2100  
QY 1221 HGKRTOGAROKFSSLYISOFLIMYSLDGGKMOITYRGNSTGTLMVFFGNVDSGIRHNIFN 1280  
Db 2101 hgkrtogarkfsslyisqflimysldgkkmqlyrgnsetglmvffgnvdsqikhnlfn 2160  
QY 1281 PPIIRYIRLPHPTHYSTINSTLPMELMGCDLNSCSMPLGNEKSAISDAQITTASSYFTNMFA 1340  
Db 2161 ppliarylrlphpthystrlstmelmgcdlncsmplgmeskaisdaqitlasyftnmfa 2220  
QY 1341 TWSPSKARLHLOGRSMAMPPOVNNPKEMLOVDFOKMTWGTGTQGVKSILTSMTVKEFL 1400  
Db 2221 twspskarlhlogrsmampvnnpkemlvqdfqtkmvtgvtcgvkslltsmvykefl 2280  
QY 1401 ISSQDGHQWTLFFONGKVKVFGQNSDFTPYVNSLDPLLTRYLRIHQSWVHQAIALRM 1460  
Db 2281 issqdgqhwtclffqngkvkvfgngdsftpyvnslldpplltrylrlhpsvvhqialrm 2340  
QY 1461 EYLGEADDLX 1471  
Db 2341 evlgeaadlly 2351

Search completed: November 17, 2001, 13:24:24  
Job time: 133 sec



**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 17, 2001, 17:25:18 ; Search time 11937.4 Seconds  
(without alignments)  
6290.589 Million cell updates/sec

Title: US-09-689-430-1  
Perfect score: 7944  
Sequence: 1 tggccaccctctctcgcgc.....agggggggggggggggggt 7944

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 10228115 segs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_est3:\*  
4: gb\_est4:\*  
5: gb\_est5:\*  
6: gb\_est6:\*  
7: gb\_est7:\*  
8: gb\_est8:\*  
9: gb\_est9:\*  
10: gb\_est10:\*  
11: gb\_est11:\*  
12: gb\_est12:\*  
13: gb\_est13:\*  
14: gb\_est14:\*  
15: gb\_est15:\*  
16: gb\_est16:\*  
17: gb\_est17:\*  
18: gb\_est18:\*  
19: gb\_est19:\*  
20: gb\_est20:\*  
21: gb\_est21:\*  
22: gb\_est22:\*  
23: gb\_est23:\*  
24: gb\_est24:\*  
25: gb\_est25:\*  
26: gb\_est26:\*  
27: gb\_est27:\*  
28: gb\_est28:\*  
29: gb\_est29:\*  
30: gb\_est30:\*  
31: gb\_est31:\*  
32: gb\_est32:\*  
33: gb\_est33:\*  
34: gb\_est34:\*  
35: gb\_est35:\*  
36: gb\_est36:\*  
37: gb\_est37:\*  
38: gb\_est38:\*  
39: gb\_est39:\*  
40: gb\_est40:\*  
41: gb\_est41:\*  
42: gb\_est42:\*  
43: gb\_est43:\*  
44: em\_esthum10:\*  
45: em\_esthum11:\*  
46: em\_esthum12:\*  
47: em\_esthum13:\*  
48: em\_esthum14:\*  
49: em\_esthum15:\*  
50: em\_esthum16:\*  
51: em\_esthum17:\*  
52: em\_esthum18:\*  
53: em\_esthum19:\*  
54: em\_esthum20:\*  
55: em\_esthum21:\*  
56: em\_esthum22:\*  
57: em\_esthum23:\*  
58: em\_esthum24:\*  
59: em\_esthum25:\*  
60: em\_esthum26:\*  
61: em\_esthum27:\*  
62: em\_esthum28:\*  
63: em\_esthum29:\*  
64: em\_esthum30:\*  
65: em\_esthum31:\*  
66: em\_esthum32:\*  
67: em\_esthum33:\*  
68: em\_esthum34:\*  
69: em\_esthum35:\*  
70: em\_esthum36:\*  
71: em\_esthum37:\*  
72: em\_esthum38:\*  
73: em\_esthum39:\*  
74: em\_esthum40:\*  
75: em\_esthum41:\*  
76: em\_esthum42:\*  
77: em\_esthum43:\*  
78: em\_esthum44:\*  
79: em\_esthum45:\*  
80: em\_esthum46:\*  
81: em\_esthum47:\*  
82: em\_esthum48:\*  
83: em\_esthum49:\*  
84: em\_esthum50:\*  
85: em\_esthum51:\*  
86: em\_esthum52:\*  
87: em\_esthum53:\*  
88: em\_esthum54:\*  
89: em\_esthum55:\*  
90: em\_esthum56:\*  
91: em\_esthum57:\*  
92: em\_esthum58:\*  
93: em\_esthum59:\*  
94: em\_esthum60:\*  
95: em\_esthum61:\*  
96: em\_esthum62:\*  
97: em\_esthum63:\*  
98: em\_esthum64:\*  
99: em\_esthum65:\*  
100: em\_esthum66:\*  
101: em\_esthum67:\*  
102: em\_esthum68:\*  
103: em\_esthum69:\*  
104: em\_esthum70:\*  
105: em\_esthum71:\*  
106: em\_esthum72:\*  
107: em\_esthum73:\*  
108: em\_esthum74:\*  
109: em\_esthum75:\*  
110: em\_esthum76:\*  
111: em\_esthum77:\*  
112: em\_esthum78:\*  
113: em\_esthum79:\*  
114: em\_esthum80:\*  
115: em\_esthum81:\*  
116: em\_esthum82:\*



117: gb\_est48:\*  
118: gb\_est49:\*  
119: gb\_est50:\*  
120: gb\_est51:\*  
121: gb\_est52:\*  
122: gb\_est53:\*  
123: gb\_est54:\*  
124: gb\_est55:\*  
125: gb\_est56:\*  
126: gb\_est57:\*  
127: gb\_est58:\*  
128: gb\_est59:\*  
129: gb\_est60:\*  
130: gb\_est61:\*  
131: gb\_est62:\*  
132: gb\_est63:\*  
133: gb\_est64:\*  
134: gb\_est65:\*  
135: gb\_est66:\*  
136: gb\_est67:\*  
137: gb\_est68:\*  
138: gb\_est69:\*  
139: gb\_est70:\*  
140: gb\_est71:\*  
141: gb\_est72:\*  
142: gb\_est73:\*  
143: gb\_est74:\*  
144: gb\_est75:\*  
145: gb\_est76:\*  
146: gb\_est77:\*  
147: gb\_est78:\*  
148: gb\_est79:\*  
149: gb\_est80:\*  
150: gb\_est81:\*  
151: gb\_est82:\*  
152: gb\_est83:\*  
153: gb\_est84:\*  
154: gb\_est85:\*  
155: gb\_est86:\*  
156: gb\_est87:\*  
157: gb\_est88:\*  
158: gb\_est89:\*  
159: gb\_est90:\*  
160: gb\_est91:\*  
161: gb\_est92:\*  
162: gb\_est93:\*  
163: gb\_est94:\*  
164: gb\_est95:\*  
165: gb\_est96:\*  
166: gb\_est97:\*  
167: gb\_est98:\*  
168: gb\_est99:\*  
169: gb\_est100:\*  
170: gb\_est101:\*  
171: gb\_est102:\*  
172: gb\_est103:\*  
173: gb\_est104:\*  
174: gb\_est105:\*  
175: gb\_est106:\*  
176: gb\_est107:\*  
177: gb\_est108:\*  
178: gb\_est109:\*  
179: gb\_est110:\*  
180: gb\_est111:\*  
181: gb\_est112:\*  
182: gb\_est113:\*  
183: gb\_est114:\*  
184: gb\_est115:\*  
185: gb\_est116:\*  
186: gb\_est117:\*  
187: gb\_est118:\*  
188: gb\_est119:\*  
189: gb\_est120:\*

190: gb\_est121:\*  
191: gb\_est122:\*  
192: gb\_est123:\*  
193: gb\_est124:\*  
194: gb\_est125:\*  
195: gb\_est126:\*  
196: gb\_est127:\*  
197: gb\_est128:\*  
198: gb\_est129:\*  
199: gb\_est130:\*  
200: gb\_est131:\*  
201: gb\_est132:\*  
202: gb\_est133:\*  
203: gb\_est134:\*  
204: gb\_est135:\*  
205: gb\_est136:\*  
206: gb\_est137:\*  
207: gb\_est138:\*  
208: gb\_est139:\*  
209: gb\_est140:\*  
210: gb\_est141:\*  
211: gb\_est142:\*  
212: gb\_est143:\*  
213: gb\_est144:\*  
214: gb\_est145:\*  
215: gb\_est146:\*  
216: gb\_est147:\*  
217: gb\_est148:\*  
218: gb\_est149:\*  
219: gb\_est150:\*  
220: gb\_est151:\*  
221: gb\_est152:\*  
222: gb\_est153:\*  
223: gb\_est154:\*  
224: gb\_est155:\*  
225: gb\_est156:\*  
226: gb\_est157:\*  
227: gb\_est158:\*  
228: gb\_est159:\*  
229: gb\_est160:\*  
230: gb\_est161:\*  
231: gb\_est162:\*  
232: gb\_est163:\*  
233: gb\_est164:\*  
234: gb\_est165:\*  
235: gb\_est166:\*  
236: gb\_est167:\*  
237: gb\_est168:\*  
238: gb\_est169:\*  
239: gb\_est170:\*  
240: gb\_est171:\*  
241: gb\_est172:\*  
242: gb\_est173:\*  
243: gb\_est174:\*  
244: gb\_est175:\*  
245: gb\_est176:\*  
246: gb\_est177:\*  
247: gb\_est178:\*  
248: gb\_est179:\*  
249: gb\_est180:\*  
250: gb\_est181:\*  
251: gb\_est182:\*  
252: gb\_est183:\*  
253: gb\_est184:\*  
254: gb\_est185:\*  
255: gb\_est186:\*  
256: gb\_est187:\*  
257: gb\_est188:\*  
258: gb\_est189:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.







QY 5920 gaagatccttgatcttcttctacgaggctcgaacgagtcgaacgaacacgacgtta 5979  
|||||  
Db 601 GAAGATCCTTGAATCTTTCTACGGGGCTGACCGCTGAGTGAAGAAACTCAGCTTAA 660  
QY 5980 gggatttggcatgatattatcaaaaagatcttcaactagatccctttaataaaa 6039  
|||||  
Db 661 GGGATTTTGGTCATGAGATATATCAAAAAGATCTTCACTAGATCTTTTAAATTAATAA 720  
QY 6040 tgaagctttaaatacaactaaagataatagaaacttgcttgacagttaccatgac 6099  
|||||  
Db 721 TGAAGTTTAAATCAATCTAAAGATATATGAGTAAACTTGCTGACAGTTTACCATGTC 780  
QY 6100 ttaataagtaggacacattatcagagcatctgtctatcttgltacatagttgctga 6159  
|||||  
Db 781 TTAATCAGTAGGACCACTATCTCAGCATCTGTATTTCTGCATCATGATGTTGCTTGA 840  
QY 6160 ctcccgctgctgtagataactagatacggagaggttaccatctggcccgagtgctga 6219  
|||||  
Db 841 CTCGCCGTCGTGATGATTAATACGATACGGAGGGGCTTACATCTGGCCCAAGTGTGCA 900  
QY 6220 atgataccgagagacccacgctcacggtccagattatcagcaataaacacgacgac 6279  
|||||  
Db 901 ATGATACCGGAGAACCCACCGCTCACCGCTCAGATTTATGCAATTAACACGACGCC 960  
QY 6280 ggaaggccgagcgacgaagtgctcgcacattatccgctccatccagctcataat 6339  
|||||  
Db 961 GGAA-GGCCGAGCGCAAGAGTGTCTGCAACTTATCCGCTCCATCCAGCTATTTAAT 1019  
QY 6340 tgttgcgggaaagctagatgaagtgtcgccagttatagttggcgaac 6390  
|||||  
Db 1020 TGTTCGGGGAGACTGATGATGATGTCGCCAGTTAATAGTTTGGCGCAAC 1070

RESULT 2

AJ281480  
LOCUS AJ281480 1004 bp mRNA EST 30-JUN-2000  
DEFINITION 4A3A-P4G8-F Anopheles gambiae Immune competent 4A3A Anopheles  
gambiae cDNA clone 4A3A-P4G8, mRNA sequence.  
ACCESSION AJ281480  
VERSION AJ281480.1 GI:6929360  
KEYWORDS EST.  
SOURCE African malaria mosquito.  
ORGANISM Anopheles gambiae.  
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoides  
; Anopheles.  
1 (bases 1 to 1004)  
Dimopoulos,G., Casavant,T.L., Chang,S., Scheetz,T., Roberts,C.,  
Donohue,M., Schultz,D., Benes,V., Bork,P., Ansoirge,W., Soares,M.B.  
and Kafatos,F.C.  
Anopheles gambiae pilot gene discovery project: identification of  
mosquito innate immunity genes from expressed sequence tags  
generated from immune-competent cell lines  
Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)

TITLE  
JOURNAL MEDLINE  
COMMENT Contact: Dimopoulos G  
Fortis C. Kafatos Laboratory  
European Molecular Biology Laboratory  
Meyerhofstrasse 1, 69117 Heidelberg, Germany.  
Location/Qualifiers  
1.1004  
/organism="Anopheles gambiae"  
/strain="4A r/r"  
/db\_xref="taxon:7165"  
/clone="4A3A-P4G8"  
/clone\_lib="Anopheles gambiae immune competent 4A3A"  
/cell\_line="Immune competent 4A3A"  
/lab\_host="E. coli DH10B"  
/note="Vector: pT73D-Pac (Pharmacia) with a modified  
polylinker. Site\_1: EcoRI; Site\_2: NotI; sequenced from  
forward priming site which reads from the 3' end of the  
cDNA. The 4A3A is a directionally cloned and normalized

BASE COUNT 252 a 262 c 244 g 244 t 2 others  
ORIGIN  
Query Match 11.6%; Score 920.8; DB 104; Length 1004;  
Best Local Similarity 98.8%; Pred. No. 1,7e-254;  
Matches 999; Conservative 1; Mismatches 4; Indels 7; Gaps 7;  
cDNA library that was constructed from the 4A3A cell line  
oligo-T primed cDNA according to: Bonaldi, Lennon & Soares  
(1996) : Normalization and Subtraction: Two Approaches To  
Facilitate Gene Discovery, Genome Research 6, 791-806."

QY 5291 aaaagccaaagacacgtaaaaaagcgcgctgtcgtgcttcttcataagctccgcccc 5350  
Db 1 AAAAGGCCAGAACCGTAAAAAGGCCCGTGTGCTGGCGTTTTCATAGAGCTCCGCCCC 60  
QY 5351 ctgacgagcatcacaaaaatgcagctcaagtcaaggttgcgaaacccgacagagctat 5410  
Db 61 CTGACGACATCACAAAATGACGCTCAAGTCAAGAGTGGCGAACCAGAGACTAT 120  
QY 5411 aaagataccaaagcgttcccccctgaaagctccctcgtgcgtctcctgttcgaacctg 5470  
Db 121 AAAGATACAGAGCGTTTCCCGCTGGAAGTCCCTGAGCGCTCTCGTTCGACCTCG 180  
QY 5471 cgcttacccgataccgtgtccgcttctcccttcgggaagcggtgacctctcaatgct 5530  
Db 181 CGCTTACCGGATACCTGTCCGCTTCTCCCTTCGGGAACGTCGGCGCTTTCATAGCT 240  
QY 5531 cagcgtgtagtactcagttcagttcagttcagttcagttcagttcagttcagttcag 5590  
Db 241 CACGCTGAGGATCTCAAGTGTGGTGTGAGTGTGCTTCCTCAAGCTGGGCTGTTACAG 300  
QY 5591 aacccccgttcaagcccgacccgctgcgtctatccgtaactatcgtctgaatccaac 5650  
Db 301 AACCCCGCTTACCGCCGACCGCGCTGCGCTTATCCGGTATTCGCTGATCCACAC 360  
QY 5651 ccgtaagaacacgactatccgcaactgacgacgacgactgttaacagattagcagcga 5710  
Db 361 CGGTAAACACACACTTATCCGCCATGCGCACAGCACGCACTGTAAACAGATTAGCAGCG 420  
QY 5711 ggtatgtaggggtgtcctacagagttcttgaagtgtgtgcttaactcggctacacaga 5770  
Db 421 GGTATGTAAGGCGGTCTACAGAGTCTTGAAGTGTGCTTAACTACGAGCTACACTGA 480  
QY 5771 ggaacatttggatctgtcgtctcgttgaaagcagttacttcgcgaaagaagttgta 5830  
Db 481 GAACAGATTGTTGATATGCGCTGTCTGTAAGCCAGTTACTTTCGGAATAAGATTGTA 540  
QY 5831 gctctgacccggcaacaacaacacgctgtgtagcgtgttcttctgttgcacagcag 5890  
Db 541 GCTCTGATCCGGCAAAACCAACCGCTGTGCGGTGTTTGTTCGCAACGACG 600  
QY 5891 agattaacgacagaaaaaagatcacaagaagatcccttgatccttccaggggtctg 5950  
Db 601 AGATTACGCGCAAAAAGATCTCAAGAGTCTTGTGATCTTTCTACGGGGTCTG 660  
QY 5951 acgctcagtggaagaaacacacgtaaggaatttggcatgagatatacaaaaaagga 6010  
Db 661 ACGCTCAGTGGAAAGAAACCTCAGTTAAGGATTTGGTCATGAGATTATCAAAAAGGA 720  
QY 6011 tcttcacctagatcctttaaataaataagatttaatacaatcaagtatatag 6070  
Db 721 TCTTCACCTAGATCTTTTAAATTAATAAGATTAAATCAATCAATGATATATG 780  
QY 6071 agtaaatgtgtcgaagaattacaatgcttaactagtaggagacattatcagagatc 6130  
Db 781 AGTAACTTGGTCTGACAGTTACCAATGCTTAATCACTGAGCA-CTATCTCAGCATCT 839  
QY 6131 gtcattctgttcaatcagatgttcgtaccccgctgctgtgataactacagatacggg 6190  
Db 840 GTCTATTTCGTTCACTCATGATGTCCTGACTCCCGCTGCTGATGAT-ACTACGATACGGG 898  
QY 6191 agggttaccatctgtgcccagtgctgcaatgataccgcgaagccacgctcaccgctc 6250



|||||  
Db 899 AGGCGCTTACATYTG-CCAGTGCTGCAATGATACCGCAGAG-CCAGCTCA-CGCGTC 955  
QY 6251 cagattatagcaataaccagccagcggaagggcgagcgcaagtg 6301  
|||||  
Db 956 CAGATT-TGAGCAATTAACCGACCGCGGA-GGCGGAGCGCAGAAGTG 1004  
RESULT 3  
AL042026/c 841 bp mRNA EST 29-FEB-2000  
LOCUS AL042026  
DEFINITION DKFZP434E111.1 434 (synonym: htes3) Homo sapiens cDNA clone  
AL042026  
ACCESSION AL042026  
VERSION AL042026.1 GI:5421372  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 841)  
AUTHORS Poustka,A., Klein,M., Mewes,H.W., Gassenhuber,J. and Wiemann,S.  
TITLE EST (Poustka, et al.)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Poustka A.J.  
Department Lehnach  
Max-Planck-Institute for Molecular Genetics  
Imestrasse 73, 14195 Berlin, Germany  
Tel: +49-30-84131623  
Fax: +49-30-84131128  
Email: poustka@mpimg-berlin-dahlem.mpg.de  
This is the 5' sequence of the clone insert  
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;  
Heidelberg/Germany) within the cDNA sequencing consortium of the  
German Genome Project.  
No sl sequence available.  
This clone (DKFZP434E111) is available at the RZPD in Berlin  
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.  
Location/Qualifiers  
1. 841  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="DKFZP434E111"  
/clone\_1lb="434 (synonym: htes3)"  
/issue\_type="testis"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/note="Vector: pSport1; Site\_1: NotI; Site\_2: SalI"  
BASE COUNT 226 a 192 c 202 g 221 t  
ORIGIN  
Query Match 10.6%; Score 841; DB 105; Length 841;  
Best Local Similarity 100.0%; Pred.No. 1.9e-231;  
Matches 841; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 6204 tggcccgatgtctgaatgataccgagagccagcgtccagcgtccagattatcagc 6263  
|||||  
Db 841 TGCGCCAGTGTGCAATGATACCGGAGACCGCTCACGGCTCCAGATTATCAGC 782  
QY 6264 aataaacccagcagcggaagggcgagcgagaagtgtctcgtcaacttatccgctc 6323  
|||||  
Db 781 AATTAACCAACCCAGCGGAGGGCCGACGCCAAGAGTGCTGCAACTTTATCCCGCTC 722  
QY 6324 catccagtcataatgtgttcgagggaagtagaagtaagtaagtcgccagttatagttt 6383  
|||||  
Db 721 CATCCAGTCTAATTAATGTGTGCGGGAAGTAGAGTAAGTTCCCGAGTTAATAGTTT 662  
QY 6384 ggcgaacgttgttcacgtctacagcagcagtcggtgtcagcgtcgttgtgtatgac 6443  
|||||

Db 661 GCCCAAGTGTGTCATGCTACAGCATGCTGTTGTCACGGCTGCTGTGGTATGCG 602  
QY 6444 ttcatcagctccggttcccaagatcaagcgagtaagtgatcccatgtgtgcaa 6503  
|||||  
Db 601 TTCAATTCAGCTCCGGTTCCCAACGATCAAGCGAGTTACATGATCCCGCATGTGTGCAA 542  
QY 6504 aaaaagcgtagctccttcgctccgcatcgtgtgtcagaagaagtgtgacgagtggt 6563  
|||||  
Db 541 AAAAGCGGTAGCTCCTTCGCTCTCGATCGTTGTGACAGTAAGTTGGCCGCGAGTGT 482  
QY 6564 atcaactatggtatgtgcaagcaactgcatlaattctctactgltcatgcatcgtlaagt 6623  
Db 481 ATCAGTCATGTTATGGCAGCAGCTGATATTTCTTACTAGTGTATGCGATCCGTAAGATG 422  
QY 6624 cttctctgtactgtgtgagtaactcaacaaagtaactctgagaatagtgtatgcggacc 6683  
Db 421 CTTTCTGTGACGTGTGAGTCAACCAAGTCATTTCTGAAATAGTGTATGCGGCGACC 362  
QY 6684 gagtgcctcttcggcggtgcaataagggatlaaccgscgcaatagcagaactttaa 6743  
|||||  
Db 361 GAGTTGCTCTTGCCCGCGCTCAATACGGGATTAATACCGCCGCAATACGAACTTTAAA 302  
QY 6744 agtgcatactatggaanaagctcttcggtcggaagaaactcgaagatctaccgctgt 6803  
|||||  
Db 301 AGTGTCATCATTTGGAANAAGCTTCTTGCGGGGCAAAACTCTCAAGATCTTACCGTGT 242  
QY 6804 gagatccagttcgatgttaaccacactcgtgcacccaactgtatcttcaagatcttactt 6863  
Db 241 GAGATCCAGTTGATGATGACCACTCTGTCACCCCACTGATGATCTTCAACATCTTTTACTTT 182  
QY 6864 caccagcgttcttcgggtgagcaaaaacaggaaggaagcgcgaagaaagggaatgaag 6923  
|||||  
Db 181 CACCAGCGTTTCTGGGTGAGCAAAAACAGCAAGCAAAAATGCCCAAAAAGGAATTAAG 122  
QY 6924 ggcgaacggaatagtgtgaactcactcctctcttcttcaatlatatgtgaagcatla 6983  
Db 121 GGGGACACGGAATGTTGAATACATACATCTCTCTTTCAATTAATTAAGAACATTTA 62  
QY 6984 tcagggtatgtctcctatgagcgagatatactattgtaattgattagaaaataacaaat 7043  
Db 61 TCAAGGTTATGTTGCTCATGAGCGGATACATTTGATTTGATAAATAAACAAAT 2  
QY 7044 a 7044  
Db 1 A 1  
RESULT 4  
AJ281449  
LOCUS AJ281449 800 bp mRNA EST 30-JUN-2000  
DEFINITION 4A3A-P4D5-F Anopheles gambiae immune competent 4A3A Anopheles  
gambiae cDNA clone 4A3A-P4D5, mRNA sequence.  
AJ281449  
ACCESSION AJ281449  
VERSION AJ281449.1 GI:6929329  
KEYWORDS EST.  
SOURCE African malaria mosquito.  
ORGANISM Anopheles gambiae  
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae  
; Anopheles.  
1 (bases 1 to 800)  
REFERENCE Dimopoulos,G., Casavant,T.L., Chang,S., Scheetz,T., Roberts,C.,  
Donohue,M., Schultz,J., Benes,V., Bork,P., Ansoerge,W., Soares,M.B.  
and Kafatos,F.C.  
Anopheles gambiae pilot gene discovery project: identification of  
mosquito innate immunity genes from expressed sequence tags  
generated from immune-competent cell lines  
Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)  
JOURNAL MEDLINE  
CONTACT: Dimopoulos G  
Fotis C. Kafatos Laboratory  
European Molecular Biology Laboratory  
Meyerhofstrasse 1, 69117 Heidelberg, Germany.



```
FEATURES
  source
    Location/Qualifiers
      1..800
        /organism="Anopheles gambiae"
        /strain="4A r/r"
        /db_xref="taxon:7165"
        /clone="4A3a-P4D5"
        /clone_lib="Anopheles gambiae Immune competent 4A3a"
        /cell_line="Immune competent 4A3a"
        /lab_host="E. coli DH10B"
        /note="Vector: p773D-Pac (Pharmacia) with a modified
polylinker. Site_1: EcoRI; Site_2: NotI; sequenced from
forward priming site which reads from the 3' end of the
cDNA. The 4A3a is a directionally cloned and normalized
cDNA library that was constructed from the 4A3a cell line
oligo-T primed cDNA according to: Bonaldo, Lennon & Soares
(1996) : Normalization and subtraction: Two approaches to
Facilitate Gene Discovery, Genome Research 6, 791-806."

BASE COUNT      203 a      206 c      198 g      193 t
ORIGIN

Query Match      9.8% Score 774.6; DB 104; Length 800;
Best Local Similarity 98.9%; Pred. No. 3.1e-212;
Matches 780; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

OY 5275 tgaagcaaaagccagcaaaagccaggaacccgttaaaagccgcttgctgagcttttc 5334
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 12 TAAGCTTGGCGGCGAAGAGCCAGAACCGTAAAGGCGCGCTTGTGCGCTTTTTC 71

OY 5335 catagctccgccccctgacgagcatcaaaaatacgaagctcaagatcagatgacga 5394
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 72 CATAGGCTCGGCCCCCTGACGAGCATCACAAAATAACGCTCAAGTCAGAGGTGGCGA 131

OY 5395 aaccgcagagactataaagatacgaagcgcttcccccctgaaagctccctgagcctc 5454
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 132 AACCCGACAGAGACTATTAAGATACCAAGCGTTTCCCTCGAAGCTCCCTGCGCTCT 191

OY 5455 cctgttccgacctgacgctacacgaatacctgtccgcttctcccttcggaaagcgtg 5514
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 192 CCGTTCGACACCTGCGGCTTACCGGATACCTGTCGCGCTTTCCTCCCTCGGGAAGCGTG 251

OY 5515 ggccttctcaatgctcaagctgtgagatctcagttcggtggtgctgcgtccaaag 5574
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 252 GCGCTTCTCATAGCTCAGCTGATGATCTCAGTTGCGGTGAGTGGTTCGCCCAAG 311

OY 5575 ctggagctgtgacgaagccccccgttcaagccgacgctgcgcttaccgtaacctat 5634
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 312 CTGGGCTGTGTGCAGAACCCCGTTACGCCGACCGCTTATCCGGTAACATAT 371

OY 5635 cgcttctgaatccaaacccggttaagaacagactatcgcacatgacgacgacacttgtaac 5694
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 372 CGCTTGATGTCNAACCCGGTAAAGACAGATTATCGCACATGGCAGCAGCCACTGTATAC 431

OY 5695 aggaatagcagagcagatgtagtcggtgcgtacagaagttcttgaagtgtgtgcctaac 5754
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 432 AGGATTATGACAGAGGAGATGTATGGCGGTCTTACAGAGTTCTTGAAGTGTGGCTTAC 491

OY 5755 taggcgtacactaagaagacagiatcttgtaatcgcgctcgtcgtgaagccagtiacctc 5814
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 492 TAGGGCTACACTAGAAACAACGATTTTGTGATCTGTGCGCTTGTGTAACGTTTACCTTC 551

OY 5815 ggaaaaaagatgtgtaactcttgaatccgcgaacaaacacacgcttgtaagcgtgtgtt 5874
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 552 GGAAGAAAGATGTGTACTCTTGTATCCGGCAAAACCAACCGGTGATAGGCGTGTGTTT 611

OY 5875 ttgtgttgcgaagcagatcagcgcaagaataaagatctcaagaagatcccttgatc 5934
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 612 TTGTGTTTGAAGAGCGAGATTTACGCGCAAGAAAAAGATCTCAAGAAAGATCTTTGATC 671

OY 5935 ttctctacgggtctcgtgcgtcagttgaaacaaactcagcttaaggagatttgatcag 5994
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 672 TTTTCTTACGGGGGTCTGACGCTCAGTGGACGAAACATCAGTTAAGGATTTTGTGATG 731
```

```
OY 5995 agatatcaaaaagatctcactagatccctttaaattaaataaagatlttaataca 6054
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 722 AGATATATCAAAAAGAGATCTTACCGTACGATTCCTTTAATTAATAAGATTTAATFCA 791

OY 6055 atctaaagt 6063
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 792 ATCTAAAGT 800

RESULT 5
LOCUS   A0081137
DEFINITION A0081137 Oncorhynchus mykiss kidney infected by infectious
hematopoietic necrosis virus Oncorhynchus mykiss cDNA clone K12,
mRNA sequence.
ACCESSION A0081137
VERSION   A0081137.1 GI:6431485
KEYWORDS EST.
SOURCE   rainbow trout.
ORGANISM Oncorhynchus mykiss
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei;
Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
REFERENCE 1 (bases 1 to 1067)
AUTHORS Sakai,M. and Kono,T.
TITLE The EST analysis of kidney and gill cells rainbow trout,
Oncorhynchus mykiss, infected with infectious hematopoietic
necrosis virus
JOURNAL Unpublished (1999)
COMMENT Contact: Masahiro Sakai
Faculty of Agriculture
Miyazaki University
1-1 nishigakuenkibanadai, Miyazaki, Miyazaki 889-2192, Japan
Email: aob208@ecc.miyazaki-u.ac.jp.

FEATURES
  source
    Location/Qualifiers
      1..1067
        /organism="Oncorhynchus mykiss"
        /db_xref="taxon:8022"
        /clone="K12"
        /clone_lib="Oncorhynchus mykiss kidney infected by
infectious hematopoietic necrosis virus"
        /tissue_type="kidney infected by infectious hematopoietic
necrosis virus"

BASE COUNT      235 a      292 c      288 g      250 t      2 others
ORIGIN

Query Match      9.7% Score 772.6; DB 107; Length 1067;
Best Local Similarity 98.6%; Pred. No. 1.3e-211;
Matches 800; Conservative 0; Mismatches 9; Indels 2; Gaps 2;

OY 5091 ccaactgataatgaatccggaacgacgagggagagagcggttgatggagcctc 5150
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 259 CCAAGTGCATTAATGAATTCGCCCAACGCGCGGAGAGCGCGTTTGCGTATTTGGCGCTC 318

OY 5151 ttccgcttcctcgtcactgactcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 5210
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 319 TTCCGCTTCCTTCGCTCACTGACTCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 378

OY 5211 agctcaactcaaaagcgtgtaatacggttatccagaatcaaggagataagcaggaagaa 5270
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 379 AGCTCAGCTCAAAAGCGGATTAATCGTTATCCACAGAAATCAGGGGATTAACACAGAAAGAA 438

OY 5271 catgtgaacaaaagcgcaagaagcgcaagaacgftaaaaagccgctgtgtgagct 5330
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 439 CATGTGACCAAAAGGCCGCAAAAGGCCAGAACCGTAAAGAGGCCGCTTGTGCGCTG 498

OY 5331 ttccatagagctccgccccctgacgagcatcaaaaatacgaactcgaatcagaagt 5390
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 499 VTTCATATGCGTCCGCCCTTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTG 558

OY 5391 ggaagaccgagagactataaagatacgaaggttcccccctggaagctccctcgtgag 5450
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```



Dh 559 GCGAAACCCGACAGACTATAAAGATACCAAGCGTTCCCTCGAAGCTCCTGTCGC 618  
QY 5451 ctctccgttccgaacctgcgcgttaccgatactctccgcttctcccttcgggaag 5510  
Db 619 CTCGCTGTTCCGACCCCTGCGCTTACCGATACCTTCGCCCTTCTCCCTGGGAAG 678  
QY 5511 cgttgcccttctcaatgcgaacgtctgaagtatcgaattcgttgtagtctgctc 5570  
Db 679 CGTGGCCCTTTCATATAGCTCAGCCCTAGTATCTAGTTCCGTGTAAGTCGTTCCCTC 738  
QY 5571 caagcttgagctgtgtgcgaacccccgcgttcagcccgacgcgtctgacctatccgtaa 5630  
Db 739 CAACCTGGGCTGTGTGACCAACCCCGTTACCGCCGCGTGGCCCTTATCCGGTAA 798  
QY 5631 ctatcgtcttgatcccaaccccggttaagaacagactatcgccactggtgcagcagcaactg 5680  
Db 799 CTATGCTTGTGAGTCCAAACCCGGTAAACAGACTTATCCCATGTCACACAGCCACTGG 858  
QY 5691 taacagattagcagaacgagatgtgtaggcgtgtctacagagttcttgaagtgtgagc 5750  
Db 859 TAACAGGATTAAGCAGACGAGGATGTAGCGGCTGCTACAGAGTTCTTGAAGTGTGCC 918  
QY 5751 taactacgcttactagaagaagacagatattggtatctgcgtctgtgaagcagttac 5810  
Db 919 TAACATGCGCTACACTAGAGAGACAGATTTGGTATCTGCGCTCTGCTGAAGCCAGTTAC 978  
QY 5811 cttcggaaaaaggtgtgtgctcttgaatccggaacaaacacgcgctgtgtgagcgtg 5870  
Db 979 CTTCGGAAGAAAGATTGTGAGCTTTGATCCGCCAAACAA-CAACGCTGTGATCGGTT 1037  
QY 5871 tttttgtttgcaagcagatctaacgcgc 5901  
Db 1038 TTTTGTGTTTGC-AGCAGCAGATTAACGCCG 1067

RESULT 6  
AL044364 954 bp mRNA EST 29-FEB-2000  
LOCUS AL044364  
DEFINITION DKFZP434C172\_s1 434 (synonym: htes3) Homo sapiens cDNA clone  
ACCESSION AL044364  
VERSION AL044364.1 GI:5432586.  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 954)  
AUTHORS Ansorge, W., Benes, V., Krieger, S., Mewes, H.W., Gassenhuber, J. and  
Wiemann, S.  
TITLE EST (Ansorge, Benes, et al.)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Ansorge W  
MIPS

Am Klopferplatz 18a D-82152 Martinsried, Germany  
This is the 3' sequence of the clone insert  
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
Research Center (DKFZ), Email: s.wiemann@dkfz-heidelberg.de;  
sequenced by EMBL (European Molecular Biology Laboratory,  
Heidelberg/Germany) within the cDNA sequencing consortium of the  
German Genome Project.  
No r1 sequence available.  
This clone (DKFZP434C172) is available at the RZPD in Berlin.  
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.  
Location/Qualifiers

FEATURES  
source  
1. 954  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="DKFZP434C172"  
/clone\_lib="434 (synonym: htes3)"  
/tissue\_type="testis"  
/dev\_stage="adult"

/lab host="DH10B"  
/note="Vector: pSport1; Site\_1: NotI; Site\_2: SalI"  
BASE COUNT 214 a 273 c 253 g 214 t  
ORIGIN

Query Match 9.7%; Score 769.2; DB 105; Length 954;  
Best Local Similarity 99.6%; Pred. No. 1.2e-210;  
Matches 771; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5095 ctccatataaagatcgcgaacgcgcgcgaggagagcggttgcgtatctggcgtctcc 5154  
Db 181 CTGATTAATGAATCGGCCCAACGGCGGAGAGCGCGTTGGGATTTGGCGCTCTCC 240  
QY 5155 gcttcctgcgtcactgactgcgtctgctgctgctgctgctgctgctgctgctgct 5214  
Db 241 GCTTCCCTCCGCTCACTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300  
QY 5215 cactcaaaagcggtaaaagctatccagaatccagggatatacaggaagaagacatg 5274  
Db 301 CACTCAAGAGCGGTAATACGATTATCCAGAAATCAGGGATTAACGAGAAAGAACATG 360  
QY 5275 tgaagcaaaagccagcaaaagccaggaacggtlaaaaggccgctgtgctgcgttttc 5334  
Db 361 TGAGCAAAAGCGCCAGCAAAAGCCAGGAACCGTAAAGCCGCGTTGCTGGCGTTTTC 420  
QY 5335 catagctccgccccctgaagcatcacaaaaatcgacgtcgaagtcaaggtgagcga 5394  
Db 421 CATAGGCTCCGCCCCCTCGAGCAATCAACAAATCGACGTCATGACAGGTGCGGA 480  
QY 5395 aaccgcaagactataaagaatccaaagggtttcccccctgaaagctccctcgtgcgtct 5454  
Db 481 AACCCGACAGGACTATAAATATACAGGCGTTTCCCTGGAAGCTCCCTGCGCGCTCT 540  
QY 5455 cctgtccgaacctgcgcgttaccgatactgttcgcgtcttctcccttcgggaagcgtg 5514  
Db 541 CTTGTTCCGACCTGCGCGCTTACCGGATACCTGTCCGCTTCTCCTTGCGGAAGCGTG 600  
QY 5515 ggcgttctcaatgctcaagctgtgaagttcgaagttcgaagttcgaagttcgaag 5574  
Db 601 GCGCTTCTCATGAGCTCACCTGAGTATCTCAGTTCGGGTGAGTTCCTCCCTCAAG 660  
QY 5575 ctgggcctgtgtgcaagaaacccccgttcagcccgacgcgcgtctatccgttaactat 5634  
Db 661 CTGGGCTGTGTGACGAACCCCGTTCAAGCCGACCGCTGCGCTTATCGGTAACATAT 720  
QY 5635 cgctctgagtcgaacccgcgtgaagaacagactatcgccactgtgcgaagcactgtgaac 5694  
Db 721 CGTCTTGAGTCCAAACCGGTAACAGACTTATTCGCACTGGCAGAGCCACTGTGAAC 780  
QY 5695 aggaattagcagaacgaggtatgtagcgtgtgtacagaagttcttgaagtgtagcctaac 5754  
Db 781 AGGATTAGCAGACGAGGTATGTAGGGCGGTACACAGAGTTCTTGAAGTGTGGCTTAAC 840  
QY 5755 taagcgtacactagaagaagatattggtatctgcgtctgctgtgaagcagttacttc 5814  
Db 841 TACGGCTACACTAGAAAGACAGATTTGGTATCTGCGCTTCTGCTGAAGCCAGATTACCTTC 900  
QY 5815 ggaaaaaagattgtgactctgtatccgcggaacaaacacgcgtgtgtagcgtg 5868  
Db 901 GGAAGAAAGAGTGTGAGCTTGTATCCGCGCAACCAACGACCGCTGTAGCGGT 954

RESULT 7  
LOCUS AU081124 1089 bp mRNA EST 15-NOV-1999  
DEFINITION AU081124 Oncorhynchus mykiss Kidney infected by infectious  
hematopoietic necrosis virus Oncorhynchus mykiss cDNA clone KG12,  
rainbow trout.  
ACCESSION AU081124  
VERSION AU081124.1 GI:6431472  
KEYWORDS EST.  
SOURCE



ORGANISM	Oncorhynchus mykiss Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus. 1 (bases 1 to 1089) Sakai,M.,and Kono,T. The EST analysis of kidney and gill cells rainbow trout, Oncorhynchus mykiss, infected with infectious hematopoietic necrosis virus Unpublished (1999)
JOURNAL	CONTACT: Masahiro Sakai Faculty of Agriculture Miyazaki University 1-1 nishi kakenkibanadai, Miyazaki, Miyazaki 889-2192, Japan Email: abd08@cc.miyazaki-u.ac.jp.
COMMENT	Location/Qualifiers 1..1089
FEATURES	/organism="Oncorhynchus mykiss" /db_xref="taxon:8022" /clone="KG12" /clone_1lb="Oncorhynchus mykiss kidney infected by infectious hematopoietic necrosis virus" /tissue_type="Kidney infected by infectious hematopoietic necrosis virus"
BASE COUNT	245 a 297 c 295 g 251 t 1 others
ORIGIN	
Query Match	9.3%; Score 741.2; DB 107; Length 1089;
Best Local Similarity	96.5%; Pred. No. 1.6e-202;
Matches 801; Conservative	0; Mismatches 23; Indels 6; Gaps 4
OY	5091 ccagctgcataatgaatcgcgcacgcgcgggaaaggcgattgcatltyggcgttc 5150
Db	260 CCAGCTGCAATTATGAATCGGGCCAAAGCGCGGGAAGAAGCGGTTCGATTGGCGCTC 319
OY	5151 ttccggttctcgtcgtactcgtacctcggtcggtcggttcgttcgttcggtcggtcggtatc 5210
Db	320 TTTCGGTTCCTCGCTACGTACGTGCGTGCTGGTTCGTTCGGCTCGCGGACGGGTATC 379
OY	5211 agctcaactaaaagcggttaacggttatccacagaatacagggtataacgcagaaaga 5270
Db	380 AGCTCACTCAAAGCGGTATACGGTTATCCACAGATCACGGGATATACCAGAAAAGA 439
OY	5271 catltgagcaaaagcgcagcaaaaggccaggaaccglaaaaaagccggtltgctgcytt 5330
Db	440 CATGTGAGCAAAAAGGCCACAAAGGCCAGAACCGTTAAAAAAGCCGCCTGTGCGCTG 499
OY	5331 ttccctaaggtctcggccccctgcagaagatatcaaaaaatcgacgtccaagttaagggtg 5390
Db	500 TTTCCATTAGGCTCCGGCCCCCTGTACGAGCATCCAAAATTCGACGGTCAAGTGAAGGTG 559
OY	5391 gcgaaccgccagacagagatataaagatacacagcggtttccccttggaagctccctgtg 5450
Db	560 GCGAAACCCGACAGAGACTATAAGATACACAGCGGTTTCCCCTGGAGACTCCCTCGTGGC 619
OY	5451 ctctccgttccgaccctcgccgttacccgagatacctgtccgcctttcccttcoggyaag 5510
Db	620 CMTCTGTTCCGACCCCTCGCGCTTACCGGATACGTCCGCTTCCTTCCTCCGGGAAG 679
OY	5511 cgtygagcttctcgaatgtcatacgcgtgtatgtatctccagttcggtgttagtgtctgc 5570
Db	680 CGTGGGCTTTTCATAGCTCAAGCTGTAGGTATCTCAGTTCGGTTAGGTTCCTTGGCTTC 739
OY	5571 caagctggagctgtgtgcagaaaccccgcttaagcccgagccgttgagcttaccggttaa 5630
Db	740 CAAGCTGGGCTGTGTACAGAACCCCCCGTTCAGCCCGACCGCTGGCCCTTATCCGGTAA 799
OY	5631 ctatcgtctctgaatccaacccggtlaagacaagcatatacgcaactlgagcaagcaacttg 5690
Db	800 CTAATGCTCTTGAAGTCCAACCCGGTAAAGACACAGACTTATCGCCACTGCGACACCACTGG 859
OY	5691 taacagattagcagagcagaggtatgtagtcggtgtctaacagagtlcttctgaag-tgltgac 5749

[illegible]



Db	72	ATCCACGAAATACAGGGGATATACCCAGGAAAGAACATGTGATGACAAAAGGCCACAGCAAAAGGC	131
Qy	5298	caggaaecggtaaaaaaggcgcgctgtctgctgctgtttccataggtctccgcccttcagca	5357
Db	132	CAGGAACCGGTAAAAAGGCCCGCTTGGTGCGNNTTTCTTCATAGGCTCCGCCCTCGAGCA	191
Qy	5358	gcatacaaaaatcgaagcccaagttaagatggtgggaaacccgaacaggtctataagta	5417
Db	192	GCATACACAAAATATGACGCTCAAGTCAAGAGAGTGGCAAAACCCGACAGGACTGTAAAGATA	251
Qy	5418	ccagagctttccccccttgaagctcccccctcgtgcgtctccctgtttccgaacctcgcgttac	5477
Db	252	CCAGGCGCTTTCCCGCTTGAAGCTCCCTCGTGGCGCTCTCTGTTCCGACACCTTCGCCCTTAC	311
Qy	5478	cgaatacctgtcgcgccttctcccttccttggaagcgtgcgccttctcctaagtctacgcgtg	5537
Db	312	CGGATACCTGTCCGCGCTTCTCCCTTTCGGGAACCGTGGCGCTTTCATATAGCTCAAGCTG	371
Qy	5538	taggtatctcaatgttcggtgtgaagtggtttcgcttccaaagctgggcgtgtgtgcacgaacccc	5597
Db	372	TAGGTATCTCAGTTCGGTGGTGAAGTCCGTTGCTGCCATCAAGCTGGGCTGTGTGCACGAACCCCC	431
Qy	5598	cgttcagcccccagagcgcgtgcgcgccttaccgcgttaactatgctttgaagtcacaaccgcgttaag	5657
Db	432	CGTTACACCCGACCGCGTCCGCTTATCCGGTACTATTCCTTGTGATGATCCAACCCCGGTAAg	491
Qy	5658	acacagactatcgcacactgtgcagcaagccacactggtlaacaagatlaagcaagcaggtatgt	5717
Db	492	ACACAGACTATTCGCACTGTGCAGCAGCCACTGTGTATACAGAGATTAGCAGAGCGAGGTATGT	551
Qy	5718	aggcggtgtctacagagatgttctgaagtgtgtgcctaaactacaggtcactacataagaagacgt	5777
Db	552	AGGGGGGTCTACAGAGAGTCTTGAAGTGGTGGCCCTACTACGTACGCTACACTAGAAAGACAGT	611
Qy	5778	attggtatctgcgcctctgtctgaagccadtttaaccttcggaaaaagatgttgaagctcttg	5837
Db	612	ATTGGTATCTGCGGCTCTGCTGTGAAGCCAGTACTTTCGAAAAAAGATTGGTAGCTCTTG	671
Qy	5838	atccgcgaacaacaacacacgcctgtgtagcg	5866
Db	672	ATCCGGCAACAAACACACGCGCTGTAGCG	700
RESULT	9		
LOCUS	A0856202/c		
DEFINITION	nbebd0001.F05f CUGI Rice BAC library (EcoRI) Oryza sativa genomic	GSS	03-NOV-1999
ACCESSION	A0856202		
VERSION	A0856202.1	GI:62065659	
KEYWORDS	GSS.		
SOURCE	Oryza sativa.		
ORGANISM	Oryza sativa.		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
AUTHORS	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;		
TITLE	Erihartoideae; Oryzeae; Oryza.		
JOURNAL	1 (bases 1 to 757)		
COMMENT	Wing, R.A. and Dean, R.A. A BAC End Sequencing Framework to Sequence the Rice Genome Unpublished (1998) Contact: Wing RA Clemson University Genomics Institute Clemson University 100 Jordan Hall, Clemson, SC 29634, USA Tel: 864 656 7288 Fax: 864 656 4293 Email: twing@clemson.edu Seq primer: TATATGACTCTCATATAGGG Class: BAC ends High quality sequence start: 29 High quality sequence stop: 232.		
FEATURES			
SOURCE	1..757	Location/Qualifiers	

BASE COUNT  
 204 a 180 c 178 g 185 t 10 others

/organism="Oryza sativa"  
 /strain="Japanica"  
 /cultivar="Nipponbare"  
 /db\_xref="taxon:4530"  
 /clone="nbe0001F05f"  
 /clone\_1lb="CUGI Rice BAC Library (EcoRI)"  
 /tissue\_type="leaf"  
 /lab\_host="E. coli DH10B"  
 /note="Vector: pBACindigo; Site\_1: EcoRI; Site\_2: EcoRI;  
 Rice is the most important food crop in the world. Half of  
 the world population, especially those inhabiting highly  
 populated areas of the humid tropics and subtropics, rely  
 on rice as their primary source of carbohydrate.  
 Monocytledonous rice is a diploid plant (2n=24) with a  
 haploid genome equivalent of 431 Mbp (Arumuganathan and  
 Earle, 1991). The relatively small genome of rice, three  
 times larger than that of Arabidopsis, makes it suitable  
 for genomic studies. In order to facilitate positional  
 cloning, physical mapping and genome sequencing of rice,  
 we have constructed a BAC library from *Oryza sativa*,  
 Nipponbare variety using EcoRI as the cloning enzyme. The  
 library contains 55,296 clones with an average insert size  
 of 121 Kb providing approximately 15 haploid genome  
 equivalents. The deep coverage allows the isolation a  
 particular sequence with a probability of 99.9 %. Three  
 high density filters, each containing 18,432 clones  
 (doubly spotted), represent the whole library for colony  
 screening and can be requested from the Clemson University  
 BAC/EST Resource Center ([www.genome.clemson.edu](http://www.genome.clemson.edu))."

Query Match	8.5%	Score	674.6	DB	234	Length	757	
Best Local Similarity	98.0%	Pred. No.	2.7e-183					
Matches 702	Conservative	0	Mismatches	12	Indels	2	Gaps	2

QY	5434	ggaagccctccctgctgctctctctctgcttcgagccctgcgcttacccgataacc-tgtccgc	5435
Db	755	GAAGGCTCTTCGCGCCCTCCCTGTTGNCAGNCNTCCGGCTTTACGGAAACNTCTTNGCC	696
QY	5493	ctttctcc-cttcggagaagctggcgcttcttcctaagcccaagcgtgtgaatctcagtt	5551
Db	695	CTTTCTCCNCTTGGAAGAGGTGCGCTTTCTCATAGCTTCACGCTGTAGATATCTCAGTT	636
QY	5552	csgttagtgcgttcgtccgctcccaagctgagctgtgtgcacgaaccccccggttcaagcccgac	5611
Db	635	CGGTGTAGGTGCTGTTCCGCTCCAACTGGCGCTGTGTGCACGAACCCCCGTTACGCCGAC	576
QY	5612	gctgcgcctatccggtgaactatcgtcttgagtcacaaacccggttaagacaagactatcgc	5671
Db	575	GCTCGCGCTTATCCGGTAATCATCTGATTCAGTCCAACCCGGTAAAGACACGACTTATGCG	516
QY	5672	cactgcgaagcagccactcgtgtaaagaattagcagagcgaggtatgtagcggtgtctaag	5731
Db	515	CACGTGCAGAGGCACACTGCTAACAGGATTAAGCAGACCGAGGTATGTAGCGGTGCTAAG	456
QY	5732	agttcttgaagtgtgcccctaactacgcgtcaactacgaagaagacagtatgttcatctgcg	5791
Db	455	AGTTCTTGAAGTGTGTGCTTACCTACCGCTACACTAGAAAGAACAGTATTTGTGTATCTCG	396
QY	5792	ctctgcctgaagccagttactaccttcggaanaaagatttgtactcttcgactcgcgcgaacaa	5851
Db	395	CTCTGCTGAAGCCAGTTACTCTTGGAANAAGTTGTAGTCTTGTATCCGGCAACAA	336
QY	5852	ccacgcgtctgtagcgtgtgttttttttgtttgctcaagcagacagattacgcgcgaanaaag	5911
Db	335	CCACCGCTGTGTAAGCGGTGTTTTTTTGTTTGTGAAGCAGCAGATTTACGGCGAGAAAAAAG	276
QY	5912	gattctcaagaagaactcttgaatcttcttcaaggggtctgtagcctatgctgagcgaagaaact	5971
Db	275	GATTCTCAAGAAGATCTTTGTGATCTTTTCTTACGGGGCTGTACCTCATGTGGAACGAAACT	216



QY 5972 cagcttaaggagcttctgctatcagatatacaaaaaagatctccactagatcccttttaa 6031  
|||||  
Db 215 CACCTTAAGGATTTTGGTCAATGATATATCAAAAAGATCTTCCACTAGATCTTTTAA 156  
QY 6032 attaaatgaagtttaatacactaagatataatagaaacttgctgcagctt 6091  
|||||  
Db 155 ATTAAATGAGATTTAATCAATCAATGATATATGAGTAACCTGCTGACAGTT 96  
QY 6092 accaatcttaatcagtagcagcactatcagcagctctctatctcgtcatcc 6147  
|||||  
Db 95 ACCAATGCTTATACGTAGGACACCTATCTCAGCATCTGTCTATTGTTTCATCC 40  
RESULT 10  
AJ281437 703 bp mRNA EST 30-JUN-2000  
LOCUS 4A3A-P4C3-F Anopheles gambiae immune competent 4A3A Anopheles  
DEFINITION gambiae cDNA clone 4A3A-P4C3, mRNA sequence.  
ACCESSION AJ281437 GI:6929317  
VERSION AJ281437.1  
KEYWORDS EST.  
SOURCE African malaria mosquito.  
ORGANISM Anopheles gambiae  
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae  
; Anopheles.  
REFERENCE 1 (bases 1 to 703)  
Dimopoulos, G., Casavant, T. L., Chang, S., Scheetz, T., Roberts, C.,  
Donohue, M., Schultz, J., Benes, V., Bork, P., Ansoorge, W., Soares, M. B.  
and Kafatos, F. C.  
Anopheles gambiae pilot gene discovery project: identification of  
mosquito innate immunity genes from expressed sequence tags  
generated from immune-competent cell lines  
Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)  
20300950  
COMMENT Contact: Dimopoulos G  
Fotis C. Kafatos Laboratory  
European Molecular Biology Laboratory  
Meyerhofstrasse 1, 69117 Heidelberg, Germany.  
FEATURES  
source  
1..703  
/organism="Anopheles gambiae"  
/strain="4A r/r"  
/db\_xref="taxon:7165"  
/clone="4A3A-P4C3"  
/clone\_1lb="Anopheles gambiae immune competent 4A3A"  
/cell\_line="Immune competent 4A3A"  
/lab\_host="E. coli DH10B"  
/note="Vector: pT73D-Pac (Pharmacia) with a modified  
polylinker; Site\_1: EcoRI; Site\_2: NotI; sequenced from  
forward priming site which reads from the 3' end of the  
cDNA. The 4A3A is a directionally cloned and normalized  
cDNA library that was constructed from the 4A3A cell line  
oligo-T primed cDNA according to: Bonaldi, Lennon & Soares  
(1996) : Normalization and Subtraction: Two approaches to  
Facilitate Gene Discovery, Genome Research 6, 791-806."

Db 144 CTCAAAGCGGTAATATACGTTATCCACAGAAATCAGGGGATTAACCCAGGAAGAACATGTG 203  
QY 5277 agcaaaagcgccaaagcgcaagcgcaagcgtaaaaagcgccgctgtgctgcttttcca 5336  
|||||  
Db 204 ACCAATAAGGCCACGAAAAGCCAGGACCGTAAANAAGGCCCGCTGCGCTTTTCCA 263  
QY 5337 taggtccgcccccttgacgagcatcaaaaaatcgacgtcaagtaaggttgcgaaa 5396  
|||||  
Db 264 TAGGCTCGCCCGCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAAGAGTGGCGAAA 323  
QY 5397 ccgcgaaggaactataaagatacagggcttcccccgtggaagctccctgtgctctcc 5456  
Db 324 CCGCAGAGACTATTAAGATACAGGCGCTTCCCTCGGAAGCTCCCTCGGCCCTCTCC 383  
QY 5457 tcttcgagccctgcgcgtctaccgataactctgcgccttctcccttcgggaaagctgagc 5516  
Db 384 TGTTCGACCCCTGCGCTTACCGGATACCTGTCCGCTTCTCCCTTCCGGGAAGCGTGGC 443  
QY 5517 gcttctcaatgctcagctgtagtatactcagctcagctgtagtgcgttcgctccaagct 5576  
Db 444 GCTTCTCATAGCTACGCTGACGCTGTAGGTATCTCAGTTCCGTTGAGGTGCTTCCCAAGCT 503  
QY 5577 gggctgtgtcagcaagaccgccgttcagccgacgcgtgccttaccggttaactatcg 5636  
Db 504 GGGCTGTGTGACGAAACCCCGCTTACACCCGACCGCGTGCCTTTATCCGTTACTATCG 563  
QY 5637 tcttaagtcacacccggtgaagacagactatgcacatgcagcagcagcactgttaacag 5696  
Db 564 TCTTAGTCCAAACCCGGTAAGACACGACTTATCCGACATGGACACACCACTGGTTACAG 623  
QY 5697 gattagcagcagagatagtagcagtgctacagagttcttgaagtggtgacctacta 5756  
Db 624 CATTAGCAGAGGAGAGTATGAGGCGGTCTACAGAGTTCTGAAGTGTGGGCTACTA 683  
QY 5757 cggctacactagaagagacag 5776  
Db 684 GGGCTACACTAGAAAGACAG 703  
RESULT 11  
AZ547271/c 887 bp DNA GSS 14-NOV-2000  
LOCUS ENMEZ76TF Entamoeba histolytica Sheared DNA Entamoeba histolytica  
DEFINITION genomic, DNA sequence.  
ACCESSION AZ547271  
VERSION AZ547271.1 GI:11169811  
KEYWORDS GSS.  
SOURCE Entamoeba histolytica.  
ORGANISM Entamoeba histolytica.  
REFERENCE 1 (bases 1 to 887)  
Eukaryota; Entamoebidae; Entamoeba.  
AUTHORS Loftus, B., Van Aken, S. and Fraser, C.  
TITLE Determination of clone end sequences from Entamoeba histolytica  
HMI:IMSS sheared DNA library  
JOURNAL unpublished (2000)  
COMMENT Contact: Brendan J Loftus  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: bjlloftus@tigr.org  
Clones are derived from the Entamoeba histolytica HMI:IMSS sheared  
DNA library  
Seq primer: M13-Forward  
Class: Shotgun  
High quality sequence start: 78  
High quality sequence stop: 780.  
Location/Qualifiers  
1..887  
/organism="Entamoeba histolytica"  
/strain="HMI:IMSS"







Dd	508	CAAGTCATTCTGAGATAATGTATGGCGGACCGAAGTTCTTCCCGCGGTCATTAAGC	567
Oy	6711	ggataataccgcgcccacatatagacaagaactttaaagtctcatcattgtgaaaaagctttcc	6770
Dd	568	GGATATATACCGCGGCCCATAGCAGAAGACTTTAAAGTGCTCATCATTTGGAAAACGTTCTTC	627
Oy	6771	ggaggaaaacctcaagagcttcacggctgttgtagatcagctcgatgtaaccacctcg	6830
Dd	628	GGGGGGAANAACCTCTCAGAGATCTACC GCCTGTGAGATTCCTCAATTCGATGATTAACCACTCG	687
Oy	6831	Tgaccaccaactgatcttcagatc	6854
Dd	688	TGCACCCACACTGATCTTCAGATC	711
RESULT	13		
LOCUS	Bf663156	757 bp	mRNA EST 19-JAN-2001
DEFINITION	963041F03.x1 C. reinhardtii CC-1690, stress condition I, normalized		
ACCESSION	Bf663156		
VERSION	Bf663156.1 GI:12253300		
KEYWORDS	EST.		
SOURCE	Chlamydomonas reinhardtii.		
ORGANISM	Chlamydomonas reinhardtii		
REFERENCE	Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;		
AUTHORS	Chlamydomonas reinhardtii		
TITLE	1 (bases 1 to 757)		
JOURNAL	Grossman,A., Davies,J., Federspiel,N., Harris,E., Hauser,C.,		
COMMENT	Lefebvre,P., McDermott,J.P., Shreger,J., Sillflow,C. and Stern,D.		
	Analyses of the Chlamydomonas reinhardtii Genome: A Model,		
	Unicellular System for Analyzing Gene Function and Regulation in		
	Vascular Plants: project phase 3		
	Unpublished (2000)		
	Contact: Charles Hauser		
	DCMB Box 91000		
	Duke University		
	Durham, NC 27708-1000		
	Tel: 919 613 8159		
	Fax: 919 613 8177		
	Email: chauser@duke.edu.		
FEATURES	Location/Qualifiers		
source	1..757		
	/organism="Chlamydomonas reinhardtii"		
	/strain="CC-1690 wild type mt+ 21gr"		
	/db_xref="taxon:3055"		
	/clone_lib="-C. reinhardtii CC-1690, Stress condition I,		
	normalized, lambda zap II"		
	/note="Vector: pBluescript II SK-; Site_1: EcoRI; Site_2:		
	XhoI; This library, constructed by John Davies and Jeffrey		
	McDemott, combines cDNAs from CC-1690 cells grown to		
	mid-log phase in TAP-N (30 min, 1hr, 4hr), TAP-S (30 min,		
	1hr, 4hr), TAP-P (4hr, 12hr, 24hr), NO3 to NH4 (30min, 1hr		
	, 4hr) and NH4 to NO3 (30min, 1hr, 4hr). PolyA mRNA was		
	purified from each sample, PCR and cDNA synthesized.		
	The cDNA was directionally cloned into lambda zap II		
	(Strategene) in the EcoRI (5') and XhoI (3') sites.		
	pBluescript II SK- plasmids were excised from the lambda		
	ZAP clones by superinfection with Exsistit (Strategene)		
	phase. The library was normalized using method 4 described		
	in Bonaldo et al (1996) Genome Research 6: 791-806."		
BASE COUNT	167 a	219 c	208 g
ORIGIN	161 t	2 others	

[illegible]

	Query Match	8.2%	Score 652.2	DB 170:	Length 757:
	Best Local Similarity	98.7%;	Pred. No. 8.2e-177;		
	Matches	668;	Conservative	0; Mismatches	8; Indels
Oy	5091	ccagctgcataatgacatcgccaacgcgcggaggagagcgcgttgcgatlyggcgctc	5150		
Ob	75	CCAGCTGCATTATGATCATCGGCCCAACGGCCGGGAGAGCGGTTCGGATTGGGCGCTC	134		

Am Klopferstr. 18a D-82152 Martinsried, Germany  
This is the 5' sequence of the clone insert  
clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de,  
sequenced by BMFZ (Biomedical Research Center at the Charité,  
Berlin/Germany) within the cDNA sequencing consortium of the German  
Genome Project.  
No sl sequence available.  
This clone (DKFZ434111.4) is available at the RZPD in Berlin.







Db 198 TTAAGAAGTCTCTTCATTTGGAANNCTGTTCTTTCGGCCGGAANCTNCNMAAGATCTTAC 139

Qy 6797 GCGTGTGAGATCCAACTTGAATTAACCCGCTGTGACCCCAACTGTATCTCAACATCT 6855

Db 138 CCGTGATGAGATCCAGTTCATGTAAACCCACTGTCGACCCCAACTGATCTTCAGCATCTT 79

Qy 6857 ttaactttaccagcgcttctcgtgtgagcaaaacacaggaagggcaaaatgcgcg 6908

Db 78 TTAAGTTTCAACGCGGTTTCNGGGGNCNAAAACAGGAGGCAAAAATGCCCTC 27

Search completed: November 18, 2001, 04:30:13  
Job time: 39895 sec







Db	244	gaagttatatgaagcgagctcttctctgcacacagatccaacttccgggtgcccgccttgg	303
Qy	300	aagcaaaaagacccctccaccccaagcgccctggaaagtctccaccccaatgctcgaagtggtg	359
Db	304	caagtgaaagtcgctgtgtgtctccgcggcgccgtcccttaacgggttaagcccttcg	363
Qy	360	tttgcaagaggaagcaaaaagccctccaccccaagccctcgagcgct---tcgagcc	416
Db	364	gtgccttgaaattccgacactgacatgcataccactttctcttctccacaggtatccgatcc	423
Qy	417	accatgcacaaatagaagctctccacactgctctctctctgctctcttgcaattctgcttaag	476
Db	424	acacatgcaaatagaagctctccacactgctctctctctgctctcttgcaattctgcttaag	483
Qy	477	gccacagaagaatactactactctggtgtgcaatgagactgtcaatggagctataatgcagaatgat	536
Db	484	gccacagaagaatactactactctggtgtgcaatgagactgtcaatggagctataatgcagaatgat	543
Qy	537	ctcgggtgagctgctcgtggtgagcgcaagaattccctctgaagtgccaaaatctttccatcc	596
Db	544	ctcgggtgagctgctcgtggtgagcgcaagaattccctctgaagtgccaaaatctttccatcc	603
Qy	597	aacacctcaagtcgtgtacaaaagaactcgtgttgtagaattcacggttaccactttccaac	656
Db	604	aacacctcaagtcgtgtacaaaagaactcgtgttgtagaattcacggttaccactttccaac	663
Qy	657	atcgcttaagccaagcgccacccctggaatggtctgtctgaagtcctacacatccagcgctgaagt	716
Db	664	atcgcttaagccaagcgccacccctggaatggtctgtctgaagtcctacacatccagcgctgaagt	723
Qy	717	tatgatacagtggttcatatacacttaagaacaatggcttcccaatccgttaagcttcatagt	776
Db	724	tatgatacagtggttcatatacacttaagaacaatggcttcccaatccgttaagcttcatagt	783
Qy	777	gttggtgtatccctcactgtaagaagctctctgtagggagctgaatatgtatgtacacacagctaa	836
Db	784	gttggtgtatccctcactgtaagaagctctctgtagggagctgaatatgtatgtacacacagctaa	843
Qy	837	agggaagaaagaatgataaagctctccctgctgtagaagccaatacatgctctgcaagctc	896
Db	844	agggaagaaagaatgataaagctctccctgctgtagaagccaatacatgctctgcaagctc	903
Qy	897	ctgaaagaagaaatggtctccaatggtcctctgacccactgtgcttaacctactatcttctc	956
Db	904	ctgaaagaagaaatggtctccaatggtcctctgacccactgtgcttaacctactatcttctc	963
Qy	957	catgttggaacctgtgtaaaagacttgaaatctcaagcctcatgtgagccctactgatatgtaga	1016
Db	964	catgttggaacctgtgtaaaagacttgaaatctcaagcctcatgtgagccctactgatatgtaga	1023
Qy	1017	gaagggaagctctgcccagaagaaagacacagacacttgccacaatatctatactatttgc	1076
Db	1024	gaagggaagctctgcccagaagaaagacacagacacttgccacaatatctatactatttgc	1083
Qy	1077	gtaatttgatgaagggaagaattggtcactctcggaaacaaagaactccctctgatatgcagatagg	1136
Db	1084	gtaatttgatgaagggaagaattggtcactctcggaaacaaagaactccctctgatatgcagatagg	1143
Qy	1137	gatgcgcgacatcgtctcggtgccttgcccaaaatgcacacagtcgaatgtgtatgtaaacagg	1196
Db	1144	gatgcgcgacatcgtctcggtgccttgcccaaaatgcacacagtcgaatgtgtatgtaaacagg	1203
Qy	1197	tctctgcgaagctcgtatgtgatagcacaggaataatcagatctatgtgcatagtatgtgaaatg	1256
Db	1204	tctctgcgaagctcgtatgtgatagcacaggaataatcagatctatgtgcatagtatgtgaaatg	1263
Qy	1257	ggacacacatcttgaagtgacactcaatatctctgaaagtgacacatatctctgttgaagac	1316
Db	1264	ggacacacatcttgaagtgacactcaatatctctgaaagtgacacatatctctgttgaagac	1323
Qy	1317	catcgccagcgctcctcttggaatatctgcacataaacttctcttactgtccaaacactctg	1376
Db	1324	catcgccagcgctcctcttggaatatctgcacataaacttctcttactgtccaaacactctg	1383
Qy	1377	atggaaccttggaacagtctctactagtcttgatcaatctctctccacacacatgatgcatg	1436
Db	1384	atggaaccttggaacagtctctactagtcttgatcaatctctctccacacacatgatgcatg	1443
Qy	1437	ggaagcttatgtaaaagtgaacagctgtctcagaagggaaccccaactacagatatgaataat	1496
Db	1444	ggaagcttatgtaaaagtgaacagctgtctcagaagggaaccccaactacagatatgaataat	1503
Qy	1497	gaagaagcgggaagactatgatgatgatctactgattcttgaaatgatatgtgtcaggttt	1556
Db	1504	gaagaagcgggaagactatgatgatgatctactgattcttgaaatgatatgtgtcaggttt	1563
Qy	1557	gatgtatgaacaactctctctctcttataccaatctcgatgtttccagaagaatccctaa	1616
Db	1564	gatgtatgaacaactctctctctcttataccaatctcgatgtttccagaagaatccctaa	1623
Qy	1617	acttggtatacatatcatgctgtctgaagaaggagactgagactatgctcccttaagctc	1676
Db	1624	acttggtatacatatcatgctgtctgaagaaggagactgagactatgctcccttaagctc	1683
Qy	1677	gcccccgatgacagaagtataataaagtcaataatttgaacaaatggtccctcagcgatgtgt	1736
Db	1684	gcccccgatgacagaagtataataaagtcaataatttgaacaaatggtccctcagcgatgtgt	1743
Qy	1737	aggaagtatacaaaaagctccgatttatggtacatacagaatgtgaacctttaagctcgtaa	1796
Db	1744	aggaagtatacaaaaagctccgatttatggtacatacagaatgtgaacctttaagctcgtaa	1803
Qy	1797	gctattcagcatgaaatcgaagaactcttggacacttaacttattgvggaagtgtgagaaca	1856
Db	1804	gctattcagcatgaaatcgaagaactcttggacacttaacttattgvggaagtgtgagaaca	1863
Qy	1857	ctgttgattatataatttaagaataacagcgacagacataatacatctaccctcaggaatc	1916
Db	1864	ctgttgattatataatttaagaataacagcgacagacataatacatctaccctcaggaatc	1923
Qy	1917	actgatactcgtccttcttattatcgaaggagatthaccacaaaggtgtgtaaacatttgaagat	1976
Db	1924	actgatactcgtccttcttattatcgaaggagatthaccacaaaggtgtgtaaacatttgaagat	1983
Qy	1977	tttccaattctgcccggaggaatatatccaataataatgacaagtgcagtgtgaagatgg	2036
Db	1984	tttccaattctgcccggaggaatatatccaataataatgacaagtgcagtgtgaagatgg	2043
Qy	2037	ccaactaaatcagatccctcgtgtgccttgaccgctaattactagtttgcgttaataatgag	2096
Db	2044	ccaactaaatcagatccctcgtgtgccttgaccgctaattactagtttgcgttaataatgag	2103
Qy	2097	agagatctagcttcaggaactatggccctctccatctcctatctgctacaaagaatctgtagat	2156
Db	2104	agagatctagcttcaggaactatggccctctccatctcctatctgctacaaagaatctgtagat	2163
Qy	2157	caaaagagaacacagataatgtgcagacaagaaggaaatgcatccgtgtttctcgatatgtat	2216
Db	2164	caaaagagaacacagataatgtgcagacaagaaggaaatgcatccgtgtttctcgatatgtat	2223
Qy	2217	gagaacccgaagctgtgtactccacagagataatacagcttcttccccaatccagctgga	2276
Db	2224	gagaacccgaagctgtgtactccacagagataatacagcttcttccccaatccagctgga	2283
Qy	2277	gtgcagcttggaagatccagagttcccaagcctccacatcagcacagatcaatggtcat	2336
Db	2284	gtgcagcttggaagatccagagttcccaagcctccacatcagcacagatcaatggtcat	2343
Qy	2337	gttttgatagtctgcagttgcagttggtttgcaatggagtggtgcaatggtataatcta	2396
Db	2344	gttttgatagtctgcagttgcagttggtttgcaatggagtggtgcaatggtataatcta	2403
Qy	2397	agcatttgagacagaactgactccctctcgtctctctctcgtatataactccataaac	2456
Db	2404	agcatttgagacagaactgactccctctcgtctctctctcgtatataactccataaac	2463



QY	2457	aaatvgltctatgaaagacacactacccctattccattctctcagagaaactggtctcatg	2516
Db	2464	aaaatgltctatgaaagacacactacccctattccattctctcagagaaactggtctcatg	2523
QY	2517	tcgatgtgaaaaacccaggttcctatgattctctgggggtgcacaaactcagacttcggaaacga	2576
Db	2524	tcgaatgaaaaacccaggttcctatgattctctgggggtgcacaaactcagacttcggaaacga	2583
QY	2577	ggcatgacccgctctacactgaaaggtttcttagtcttgaaagaacacactggtctattctagag	2636
Db	2584	ggcatgacccgctctacactgaaaggtttcttagtcttgaaagaacacactggtctattctagag	2643
QY	2637	gacagcttatgaaagtattttcagatcattctgtctgagttaaacaatgccaattgaaaccaaga	2696
Db	2644	gacagcttatgaaagtattttcagatcattctgtctgagttaaacaatgccaattgaaaccaaga	2703
QY	2697	agctctccccaagaattccaagacacccctagcacttagcgaaagccaatttaatgycaccoca	2756
Db	2704	agctctccccaagaattccaagacacccctagcacttagcgaaagccaatttaatgycaccoca	2709
QY	2757	ccagctctgaaacgccaatcaacggaataactcgttactactcttcagtcagatcaaaag	2816
Db	2710	-----gaaatnaactcgttactactcttcagtcagatcaaaag	2745
QY	2817	gaaattgacctatgatataatccatatagtttgaaatgaaagaaggaagattttgacatttat	2876
Db	2746	gaaattgacctatgatatgataccatatagtttgaaatgaaagaaggaagattttgacatttat	2805
QY	2877	gatbagagatgaaatccagagccccccgacgcttccaagaagaacaacgcagacattattat	2936
Db	2806	gatbagagatgaaatccagagccccccgacgcttccaagaagaacaacgcagacattattat	2865
QY	2937	gctcgaatgagaagagctcttggggattctgggaatgaatgaactcccaacatggtcttaagaac	2996
Db	2866	gctcgaatgagaagagctcttggggattctgggaatgaatgaactcccaacatggtcttaagaac	2925
QY	2997	aggagctcaaggttgacagtgctccctccagctccaagaagaatgtgtttccagaagaattacgat	3056
Db	2926	aggagctcaaggttgacagtgctccctccagctccaagaagaatgtgtttccagaagaattacgat	2985
QY	3057	ggctcccttctactcagcccttataccgltgagaactaaatgaaactttggagctccctggg	3116
Db	2986	ggctcccttctactcagcccttataccgltgagaactaaatgaaactttggagctccctggg	3045
QY	3117	ccatataaagacgaagaagtgtgagaataatcatcagtgtaactttcaagaacacagccctc	3176
Db	3046	ccatataaagacgaagaagtgtgagaataatcatcagtgtaactttcaagaacacagccctc	3105
QY	3177	cgctccattctccttattcttaagccttattcttcttaatgaggaagatcagagggcaaggga	3236
Db	3106	cgctccattctccttattcttaagccttattcttcttaatgaggaagatcagagggcaaggga	3165
QY	3237	gaaacttgaaaaaactttgtcagaagcctaataatgaaacaaacttacttttgaaaagtcca	3296
Db	3166	gaaacttgaaaaaactttgtcagaagcctaataatgaaacaaacttacttttgaaaagtcca	3225
QY	3297	catcataatgaccccaactaaagaatgagtttgcactgcacaaagcttggtgcttattctcgat	3356
Db	3226	catcataatgaccccaactaaagaatgagtttgcactgcacaaagcttggtgcttattctcgat	3285
QY	3357	gttgaacttgaaaaaagaatgtgcactcagagcctgatttgaaacccctctggtctgcacact	3416
Db	3286	gttgaacttgaaaaaagaatgtgcactcagagcctgatttgaaacccctctggtctgcacact	3345
QY	3417	aaacacactgaaacctgtcctatgaggaacagatgaaagatcagaagatttgacctgttttc	3476
Db	3346	aaacacactgaaacctgtcctatgaggaacagatgaaagatcagaagatttgacctgttttc	3405
QY	3477	acacactcttgaagagaccaaagaactgtgtaacttctacttgaaataatgtgaagaagaactcgag	3536
Db	3406	acacactcttgaagagaccaaagaactgtgtaacttctacttgaaataatgtgaagaagaactcgag	3465

QY	3537	gccccgcaatc	atccagatggaagacccactttaagagaat	atgcgtccatgca	3536
Db	3466	gccccctgcaatc	acagatggaagatccaccctttaagagaat	atgcgtccatgca	3535
QY	3597	atcaatgctcaat	aaatggatataccctggcttgatggtccagatataa	gggt	3656
Db	3526	atcaatgctcaat	aaatggatataccctggcttgatggtccagatataa	gggt	3585
QY	3657	cgatgtaatc	tgctcgaacatgagcaatgaaacatcatctatcatc	atccatgga	3716
Db	3586	cgatgtaatc	tgctcgaacatgagcaatgaaacatcatctatcatc	atccatgga	3645
QY	3717	catgtgtc	actatgcaaaaaaaggaggtataaaatgcgcgttaacatctc	tcca	3776
Db	3646	catgtgtc	actatgcaaaaaaaggaggtataaaatgcgcgttaacatctc	tcca	3705
QY	3777	gggtgttttc	gacaagatgtaacatcccaagcgcgatgttgccgggttgaaatgc		3836
Db	3706	gggtgttttc	gacaagatgtaacatcccaagcgcgatgttgccgggttgaaatgc		3765
QY	3837	cttatggtcgag	catctacatgcttggatgtagcacaactttctgtgtacagcaataag		3886
Db	3766	cttatggtcgag	catctacatgcttggatgtagcacaactttctgtgtgtacagcaataag		3825
QY	3897	tgtaagactcccc	gggaatgagctctgcgaacatttagagattttccagattacagcttca		3956
Db	3826	tgtaagactcccc	gggaatgagctctgcgaacatttagagattttccagattacagcttca		3885
QY	3957	ggacaata	atgacagtggcccccaagcctgcgcagactcatctatctccgatac	taacat	4016
Db	3886	ggacaata	atgacagtggcccccaagcctgcgcagactcatctatctccgatac	taacat	3945
QY	4017	ggccggagac	cccaaggagccctttcttgatataagtgatctgtttggcccaatggt		4076
Db	3946	ggccggagac	cccaaggagccctttcttgatataagtgatctgtttggcccaatggt		4005
QY	4077	attcaacg	catcaagaccagggctgcgcgttcagaagtcttcacagcctctacatcc	ctcag	4136
Db	4006	attcaacg	catcaagaccaggggtgcgcgttcagaagtcttcacagcctctacatcc	ctcag	4085
QY	4137	tttatcatcat	gtatagctcttgatggagaagtgtagaacattacggaggaaatccact		4196
Db	4066	tttatcatcat	gtatagctcttgatggagaagtgtagaacattacggaggaaatccact		4125
QY	4197	ggaaccta	atgagctctcttgccaatgtggaatcgcgggataaacaacatactttt		4256
Db	4126	ggaaccta	atgagctctctcttgccaatgtggaatcgcgggataaacaacatactttt		4185
QY	4257	aaacctccaat	tatgtctgatacatccgtttgcaccaactcatattacatgcacgc		4316
Db	4186	aaacctccaat	tatgtctgatacatccgtttgcaccaactcatattacatgcacgc		4245
QY	4317	acctctgc	catgagttgagtgaggctgttgatttaaatagttagtagaataccatttggaatg		4376
Db	4246	acctctgc	catgagttgagtgaggctgttgatttaaatagttagtagaataccatttggaatg		4305
QY	4377	gagagtaaa	gcaatatacagaatgcacagattactgcttcatccataattacataatg		4436
Db	4306	gagagtaaa	gcaatatacagaatgcacagattactgcttcatccataattacataatg		4365
QY	4437	ggcaccc	gtgctctcttcaaaaagcttcgaactcaacctccaaggaggaatgagctggaga		4486
Db	4366	ggcaccc	gtgctctctcttcaaaaagcttcgaactcaacctccaaggaggaatgagctggaga		4425
QY	4497	ccctaaag	tgaataatccaaaaggctgcgcgaatggaattccagaatgacaatgaatgc		4556
Db	4426	ccctaaag	tgaataatccaaaaggctgcgcgaatggaattccagaatgacaatgaatgc		4485
QY	4557	acagagat	aaactactcaaggagtataaactctgcttaccagcaatgtatgtgaaagatgc		4616
Db	4486	acagagat	aaactactcaaggagtataaactctgcttaccagcaatgtatgtgaaagatgc		4545
QY	4617	ctcaatcc	acagatcaaaatgagctgaatgtgaactctcttttttcagaatgtgacaatga		4676



[illegible]



```
QY 6837 caactgattctcagcatcttacttccaccagcttctgtgtgagcaaaaacaggaag 6896
|||||
Db 6533 caactgattctcagcatcttacttccaccagcttctgtgtgagcaaaaacaggaag 6592
|||||
QY 6887 gcaaaatgcgcgaacaaaggaataagggcgacacggaatgcttgaatactactctt 6936
|||||
Db 6593 gcaaaatgcgcgaacaaaggaataagggcgacacggaatgcttgaatactactctt 6652
|||||
QY 6957 ccttttcaataatttgaacatttaccaggttatgtctctctgtgagcgatcatatt 7016
|||||
Db 6653 ccttttcaataatttgaacatttaccaggttatgtctctctgtgagcgatcatatt 6712
|||||
QY 7017 tgaatgtatttagaanaaaataacaataaggggttcgcgcgacattcccccgaatgctc 7076
|||||
Db 6713 tgaatgtatttagaanaaaataacaataaggggttcgcgcgacattcccccgaatgctc 6772
|||||
QY 7077 accctgaagcttaagaacacattatcatatgacataaactataaaaatagcgatcac 7136
|||||
Db 6773 accctgaagcttaagaacacattatcatatgacataaactataaaaatagcgatcac 6832
|||||
QY 7137 gaggccttctgtctcgcgcttccgtgtgacggtgtaaaactctgacacatgacgct 7196
|||||
Db 6833 gaggccttctgtctcgcgcttccgtgtgacggtgtaaaactctgacacatgacgct 6892
|||||
QY 7197 ccgcgagacgcgtcacacgcttctcgttaagcgatgcccgcgagacagacgcgcttaagg 7256
|||||
Db 6893 ccgcgagacgcgtcacacgcttctcgttaagcgatgcccgcgagacagacgcgcttaagg 6952
|||||
QY 7257 cgcgtcacgcggtgtgtcgcggtgtcggggctggtcctaactatgcygcacatcagcgagat 7316
|||||
Db 6953 cgcgtcacgcggtgtgtcgcggtgtcggggctggtcctaactatgcygcacatcagcgagat 7012
|||||
QY 7317 tgtactaagagtgcacatatgcsgtgtgaaataccgcacagatgcygtaaggagaataa 7376
|||||
Db 7013 tgtactaagagtgcacatatgcsgtgtgaaataccgcacagatgcygtaaggagaataa 7032
|||||
QY 7377 ccgcacatcagaagaattgttaaacgcttaatatgtgttaaatcgcgttaaatgtgttaa 7436
|||||
Db 7033 -----aaattgttaaacgcttaatatgtgttaaatcgcgttaaatgtgttaa 7082
|||||
QY 7437 atcagctcaattttttaaccatagagccggaatcgcgcaaaatcccttaataatcaaaagaa 7496
|||||
Db 7083 atcagctcaattttttaaccatagagccggaatcgcgcaaaatcccttaataatcaaaagaa 7142
|||||
QY 7497 tagcccgagataggtgtgaggtgtgtcccgatttggaacaaagatccacttttaaaagac 7556
|||||
Db 7143 tagcccgagataggtgtgaggtgtgtcccgatttggaacaaagatccacttttaaaagac 7202
|||||
QY 7557 gtggaactccaaagtcgtaagggcgaaacacgctcatcagggcgatgccccactgtgaa 7616
|||||
Db 7203 gtggaactccaaagtcgtaagggcgaaacacgctcatcagggcgatgccccactgtgaa 7252
|||||
QY 7617 ccatacctcaatcaagttttttgggttcgaggtgcggttaagaacataaalcgaaacct 7676
|||||
Db 7263 ccatacctcaatcaagttttttgggttcgaggtgcggttaagaacataaalcgaaacct 7322
|||||
QY 7677 aaaggggagcccgcatatttagagcttgacggggaagccggcggaacgtggtcggaagaaagaa 7736
|||||
Db 7223 aaaggggagcccgcatatttagagcttgacggggaagccggcggaacgtggtcggaagaaagaa 7382
|||||
QY 7737 gggaagaagaagcgaagaagcgagcgctagggcgcttgcaaatgtagcggtacacgtcgcc 7796
|||||
Db 7283 gggaagaagaagcgaagaagcgagcgctagggcgcttgcaaatgtagcggtacacgtcgcc 7442
|||||
QY 7797 gtaaccacacacccgcgcgcttaattgctgcgctacagggcgcgct 7842
|||||
Db 7443 gtaaccacacacccgcgcgcttaattgctgcgctacagggcgcgct 7488
|||||
```

```
; Patent NO. 6221349
; GENERAL INFORMATION:
; APPLICANT: Couro, Linda B.
; APPLICANT: Colosi, Peter C.
; TITLE OF INVENTION: ADENO-ASSOCIATED VECTORS FOR EXPRESSION OF FACTOR VIII
; TITLE OF INVENTION: BY TARGET
; FILE REFERENCE: AVIGEN-03743
; CURRENT APPLICATION NUMBER: 60/125,974
; EARLIER APPLICATION NUMBER: 60/125,974
; EARLIER FILING DATE: 1999-03-24
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ. ID NOS.: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 11933
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-09-364-862-13
```

```
Query Match 83.2% Score 6608.4; DB 4; Length 11933;
Best Local Similarity 91.9%; Pred. No. 0;
Matches 7211; Conservative 0; Mismatches 256; Indels 379; Gaps 7;

QY 15 ctgcgcctcctcgtcctcactgagagccggcgacaaagctcgcgacgcccggctt 74
|||||
Db 4 ctgcgcctcctcgtcctcactgagagccggcgacaaagccggcgacgcccggctt 63
|||||
QY 75 gccgcggcgccctcagtgagcgagcgagcgcgagagggagtggtgccaatccatcatc 134
|||||
Db 64 gttccgcggcgccctcagtgagcgagcgagcgcgagagggagtggtgccaatccatcatc 123
|||||
QY 135 aggggttctc-----cagatctcttctcagtaagtaaacagtaacatgacatt 180
|||||
Db 124 aggggttctcctcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 183
|||||
QY 181 acccggtgtcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 240
|||||
Db 184 ttcttaataaccatccagagagatgttgttcttaataaccatccagagatgttgttaa 243
|||||
QY 241 gctggggtgtgcataagccatgacgcga-tgcggatcctagtggttttccaagagg 299
|||||
Db 244 gaagtaattagagcgaggtcttctgcacacagatccattccgggtgcccgcctagg 303
|||||
QY 300 aagcaaaagcctcctccacccagcgctggaatgtttccacccaatgtcgaagcagttggt 359
|||||
Db 304 caggtlaagtgccgtgtgtgtgttcccgcggcctggtccttcaagggttaaggccctgc 363
|||||
QY 360 ttccaagaggaagcaaaagcctcctccacccagcgctgagcctgagagct---tcgacc 416
|||||
Db 364 gtgccttggaattactgacacgacatccatcttttcttctccaaaggtatgatcc 423
|||||
QY 417 accatgcaaatagagctcctcaactgcttcttctgtgcttcttgagatctgcttaagt 476
|||||
Db 424 accatgcaaatagagctcctcaactgcttcttctgtgcttcttgagatctgcttaagt 483
|||||
QY 477 gccacccaagaatactactcctgtgtgcagtggaactgttcagtggaactatgtcaagatgat 536
|||||
Db 484 gccacccaagaatactactcctgtgtgcagtggaactgttcagtggaactatgtcaagatgat 543
|||||
QY 537 ctgcgtgagctcctgtgagcgaagaattcctcctagagtgccaatacttttccattc 596
|||||
Db 544 ctgcgtgagctcctgtgagcgaagaattcctcctagagtgccaatacttttccattc 603
|||||
QY 597 aacacctcagtcgtgtacaaaagacctctgtttgtaaatcaggttccacttttcaac 656
|||||
Db 604 aacacctcagtcgtgtacaaaagacctctgtttgtaaatcaggttccacttttcaac 663
|||||
```



QY	657	atcgcctaagccaaagcgcaacccttgatggtctgtctaagtcttaacalccagcgctgaagttc	716
Db	664	atcgctaaagccaaagcgccacccttgatggtgctctgtctaaagtcttaacatccagctgaagttc	723
QY	717	tatgatacaagttgtcaattacaacttaagaacaatggtctcccaactctgtccaagcttcacagttc	776
Db	724	tatgatacaagttgtcaattacaacttaagaacaatggtctcccaactctgtccaagcttcacagttc	783
QY	777	gttctgtctatcctaacttggaagaagctctctgaaagagcttgataatgatatgataccaagcttaaca	836
Db	784	gttctgtctatcctaacttggaagaagctctctgaaagagcttgataatgatatgataccaagcttaaca	843
QY	837	agggagaaagaagaatgtataaagtctctccctcggttgagaaagccaataatgtcttgcaagttc	896
Db	844	agggagaaagaagaatgtataaagtctctccctcggttgagaaagccaataatgtcttgcaagttc	903
QY	897	ctgaagaagaatgtctccaatgtgctctctgaccaccaatgtgaccttaactactcaatactctctc	956
Db	904	ctgaagaagaatgtctccaatgtgctctctgaccaccaatgtgaccttaactactcaatactctctc	963
QY	957	catgttgaaaccttggtaaaagaacttgaatctcagacctcaatlttgagacctcaactagtatgtaga	1016
Db	964	catgttgaaaccttggtaaaagaacttgaatctcagacctcaatlttgagacctcaactagtatgtaga	1023
QY	1017	gaaggagatctggccaaagaaagacaagaaaccttgcaacaattataactactctttgtct	1076
Db	1024	gaaggagatctggccaaagaaagacaagaaaccttgcaacaattataactactctttgtct	1083
QY	1077	gatatgttgaagaagggaaaatttggaatctcgaatcgaaacaaagaactctcttgatgcagaatag	1136
Db	1084	gatatgttgaagaagggaaaatttggaatctcgaatcgaaacaaagaactctcttgatgcagaatag	1143
QY	1137	gatctgcgaatctctctcggtgaccttgccctcaaaaatgcacacagttcaatgtgtatgtataaacag	1196
Db	1144	gatctgcgaatctctctcggtgaccttgccctcaaaaatgcacacagttcaatgtgtatgtataaacag	1203
QY	1197	tctctgcgaagttctgatattggaatgcccacaggaataatcagttactatgtgcaltgtatggaaatg	1256
Db	1204	tctctctcgaagttctgatattggaatgcccacaggaataatcagttactatgtgcaltgtatggaaatg	1263
QY	1257	ggcacaacatctctgaatgtgacataatcttctcgaagaagtcacaataatctctgtatgaagaac	1316
Db	1264	ggcacaacatctctgaatgtgacataatcttctcgaagaagtcacaataatctctgtatgaagaac	1323
QY	1317	catctgcgaagcgctctctctgtgaaatctctgcacaataaactctctactgtctcaacaacactctg	1376
Db	1324	catctgcgaagcgctctctctgtgaaatctctgcacaataaactctctactgtctcaacaacactctg	1383
QY	1377	atggaccttgagacagttcttactgttttgttcatactctctcccaacaablatgtgcagt	1436
Db	1384	atggaccttgagacagttcttactgttttgttcatactctctcccaacaablatgtgcagt	1443
QY	1437	gaagcttattgtcaaaagtatgacagctgtgccagaggaaccccacaactacgaatgaaaaataat	1496
Db	1444	gaagcttattgtcaaaagtatgacagctgtgccagaggaaccccacaactacgaatgaaaaataat	1503
QY	1497	gaagaagcggaagaactatgatagtatcttaactacatcttgaatggatgtgtgtcaagttt	1556
Db	1504	gaagaagcggaagaactatgatagtatcttaactacatcttgaatggatgtgtgtcaagttt	1563
QY	1557	gatgatgacaacactctccctcttacttataccaattctgcagttgtgcagaagaacatccctaaa	1616
Db	1564	gatgatgacaacactctccctcttacttataccaattctgcagttgtgcagaagaacatccctaaa	1623
QY	1617	acttggttatcataatctgtctgtcgtgaagaggaagactctggactatgtctcccttaagttctc	1676
Db	1624	acttggttatcataatctgtctgtcgtgaagaggaagactctggactatgtctcccttaagttctc	1683
QY	1677	ggccccgatgatgacgaaggttatataaagacaaatatttgatacaaatggtcccttcagcggaattgt	1736
Db	1684	ggccccgatgatgacgaaggttatataaagacaaatatttgatacaaatggtcccttcagcggaattgt	1743
QY	1737	aggaaatatacaaaaagtctcgatttatgtgcataccaatgataaacctttaagactctgtga	1796

Db	1744	aggaagatcaaaaaagctccgattttagtgatcacagatgaaacctttagactgcyaa	1803
Qy	1797	gcatctcacatgaaatcagaaatcttvggaccttacttatavggggaagttagacacaa	1856
Db	1804	gcatctcagcatgaaatcagaaatcttvggaccttacttatttaggggaagttagacacaa	1863
Qy	1857	ctgttgattatatttaagaatcaagaacagacacatataacatctaccctcaagaaatc	1916
Db	1864	ctgttgattatatttaagaatcaagaacagacacatataacatctaccctcaagaaatc	1923
Qy	1917	actgagtctcgcccttcttgattccaaaggaaattaccaaaaaggtgaacaattttagaat	1976
Db	1924	actgagtctcgcccttcttgattccaaaggaaattaccaaaaaggtgaacaacattttagaat	1983
Qy	1977	tttcccaatctgcacagagaataattccaataataatgacagctgactgtagaacatggtg	2036
Db	1984	tttcccaatctgcacagagaataataattccaataataatgacagctgactgtagaacatggtg	2043
Qy	2037	ccaactaaatcagatctccgtgtgcctgacacgctaatactactagttcttcgttaatavgag	2096
Db	2044	ccaactaaatcagatctccgtgtgcctgacacgctaatactactagttcttcgttaatavgag	2103
Qy	2097	agagatctagcttccaggaatctatttggccctctccatctctgtctaaagaatctgtagat	2156
Db	2104	agagatctagcttccaggaatctatttggccctctccatctctgtctaaagaatctgtagat	2163
Qy	2157	caaagaggaacacagaataatgtccacaagaagaaatgtcaatcctgtttctgtatttat	2216
Db	2164	caaagaggaacacagaataatgtccacaagaagaaatgtcaatcctgtttctgtatttat	2223
Qy	2217	gagaaacgaagctgtaacctcacaagaataatacaagccttctcccaatccagctgga	2276
Db	2224	gagaaacgaagctgtaacctcacaagaataatacaagccttctcccaatccagctgga	2283
Qy	2277	gtgcagcttggaagatccgaagttccaagcctccaacatcatgacaagcatcaatgtgcat	2336
Db	2284	gtgcagcttggaagatccgaagttccaagcctccaacatcatgacaagcatcaatgtgcat	2343
Qy	2337	gtttttgatagtttgcagttgtcgaatttgcataagtgagtggaactgtgatacttca	2396
Db	2344	gtttttgatagtttgcagttgtcgaatttgcataagtgagtggaactgtgatacttca	2403
Qy	2397	agcatgtgagcacagaactgacttcccttcctgtctctctctctctgataataccctcaaacac	2456
Db	2404	agcatgtgagcacagaactgacttcccttcctgtctctctctctctgataataccctcaaacac	2463
Qy	2457	aaaaatgcttatagaagaacacacatccatctccatcttccagggaaacgtgtctatg	2516
Db	2464	aaaaatgcttatagaagaacacacatccatctccatcttccagggaaacgtgtctatg	2523
Qy	2517	tccagtgaaacacacagctgtatgagttccgggtgcccacaactccagacttccggaaacga	2576
Db	2524	tccagtgaaacacacagctgtatgagttccgggtgcccacaactccagacttccggaaacga	2583
Qy	2577	ggcatgacgccttactgtaaggttctctagtttgacaagaacacgtgattattacagag	2636
Db	2584	ggcatgacgccttactgtaaggttctctagtttgacaagaacacgtgattattacagag	2643
Qy	2637	gacagttatgaaagataattcagacatactgtcgtgataaanaacatgccaattgaaccaaga	2696
Db	2644	gacagttatgaaagataattcagacatactgtcgtgataaanaacatgccaattgaaccaaga	2703
Qy	2697	agcttctcccaagaattcaagaacacccctgcataggaacgaacatttaatgccaaccca	2756
Db	2704	agcttctcccaagaattcaagaacacccctgcataggaacgaacatttaatgccaaccca	2709
Qy	2757	ccagctctgaaacgcacatcaacgggaataaactcgtactactctcagtcagatcaagaag	2816
Db	2710	ccagctctgaaacgcacatcaacgggaataaactcgtactactctcagtcagatcaagaag	2745
Qy	2817	gaaatgtcatatgatataccatcatcagttgaatgtgaagaaggaagatttgaacatttat	2876



Dh	2746	gaatttgaactatgatgatataccaataatgaatttgaagaaggaaggaattttgacattat	2805
Qy	2877	gataaggataatgaataatcagagcccccgagctttccaagaagaacaacgaacattattat	2936
Dh	2806	gataaggatgtaaaataatcaagagcccccgagctttccaagaagaacaacgacctattatc	2865
Qy	2937	gctgcagatgtagagagctctcvggattatgtagatgagtaagctcccacalgttctaagaac	2996
Dh	2866	gctgcagatgtagagagagctctcvggattatgtagatgtaagctcccacalgttctaagaac	2925
Qy	2997	agggctcaagagtgaggagtgcccttcagcttcaagaagaagtgtttcccaagaattacagat	3056
Dh	2926	agggctcaagagtgaggagtgcccttcagcttcaagaagaagtgtttcccaagaattacagat	2985
Qy	3057	ggccctcttactcagcccttataccagttagaagaactcaatgtaacatttggactccggg	3116
Dh	2986	ggccctcttactcagcccttataccagttagaagaactcaatgtaacatttggactccggg	3045
Qy	3117	ccatatataagagcagaagaatttgaagaataatcatatgtaactttcagaatatcagccct	3176
Dh	3046	ccatatataagagcagaagaatttgaagaataatcatatgtaactttcagaatatcagccct	3105
Qy	3177	cgctccattcccttcatcttcttaagcccttatttttttgggaagatccaaggccaaggcca	3236
Dh	3106	cgctccattcccttcatcttcttaagcccttatttttttgggaagatccaaggccaaggcca	3165
Qy	3237	gaacctagaaaaaaacttgytcaagcccttaatgtaaaccaaaacttacttttggaaagtcca	3296
Dh	3166	gaacctagaaaaaaacttgytcaagcccttaatgtaaaccaaaacttacttttggaaagtcca	3225
Qy	3297	catcataatgagccccaactaaagatgagtttgaatgcaagccctgggttatttctctgat	3356
Dh	3226	catcataatgagccccaactaaagatgagtttgaatgcaagccctgggttatttctctgat	3285
Qy	3357	gttgacctggaaaaaagatgtagcactcagccctgatattggaacccctctgtctgccaact	3416
Dh	3286	gttgacctggaaaaaagatgtagcactcagccctgatattggaacccctctgtctgccaact	3345
Qy	3417	aacacactggaacccctctctctctgtagaagaacagtgaacagatcagaagatttggctgttttc	3476
Dh	3346	aacacactggaacccctctctctctgtagaagaacagtgaacagatcagaagatttggctgttttc	3405
Qy	3477	accaatctttagagacccaanaagctggttaacttccactggaanaatataggaagaactgtagg	3536
Dh	3406	accaatctttagagacccaanaagctggttaacttccactggaanaatataggaagaactgtagg	3465
Qy	3537	gtctccctgcgaataatccaagatgtagaagatcccaacttttaagagaattatcgcttccatgca	3596
Dh	3466	gtctccctgcgaataatccaagatgtagaagatcccaacttttaagagaattatcgcttccatgca	3525
Qy	3597	atcaaatggtgcataataatgtagatacacaactgtgtctaagtaatggtgcagaatccaagaatt	3656
Dh	3526	atcaaatggtgcataataatgtagatacacaactgtgtctaagtaatggtgcagaatccaagaatt	3585
Qy	3657	cgatgtagatctgtctcagcatgtaggcaagcaaatgtaaaaacttccattcatcttaagtgga	3716
Dh	3586	cgatgtagatctgtctcagcatgtaggcaagcaaatgtaaaaacttccattcatcttaagtgga	3645
Qy	3717	catgtgttcaactgtatcagaaaaaaagagagatataaaatgtagcaactgtgtacaaatcttatcca	3776
Dh	3646	catgtgttcaactgtatcagaaaaaaagagagatataaaatgtagcaactgtgtacaaatcttatcca	3705
Qy	3777	gggtgtttttagaagcagtggaagaatggttaccataccaaagcttggaaatttggccgggtggaatgc	3836
Dh	3706	gggtgtttttagaagcagtggaagaatggttaccataccaaagcttggaaatttggccgggtggaatgc	3765
Qy	3837	cttaatgtagagatcatcatagctgtctggaatggaagacaacttttctgtgttaagaagaataag	3896
Dh	3766	cttaatgtagagatcatcatagctgtctggaatggaagacaacttttctgtgttaagaagaataag	3825
Qy	3837	tgctaaagctccctctgtagaattgcttcttggaacaattagaaatttccaatttaacagcttca	3956
Dh	3826	tgctaaagctccctctgtagaattgcttcttggaacaattagaaatttccaatttaacagcttca	3885

[illegible]



[illegible][illegible]



```

Db 6893 ccggagagacggtccacagctgtctcgttaagcggatgcccggagcagacagaccccgtaagg 6552
Qy 7257 cgcgtccagcgggtgtctcgcgagggtctcgagggtcgttcaactatgcygcatacgaagcagat 7316
Db 6953 cgcgtccagcgggtgtctcgcgagggtcgttcaactatgcygcatacgaagcagat 7012
Qy 7317 tgtactgagaaatgtcacatatagtcgtgtgtgnaataccgcacagatgcgttaagagaaata 7376
Db 7013 tgtactgagaaatgtcacata----- 7032
Qy 7377 ccgcatacagaaatgttaaacgcttaatatgtgttaaaatctcgtaaatcttctgttaa 7436
Db 7033 -----aaatgttaaacgcttaatatgtgttaaaatctcgtaaatcttctgttaa 7082
Qy 7437 atcagctcatcttcttaaccataagcgcgaataatcgcgaanaatcccttaataatcaagaagaa 7496
Db 7083 atcagctcatcttcttaaccataagcgcgaataatcgcgaanaatcccttaataatcaagaagaa 7142
Qy 7497 tagagccgagataagggttggatgtgtgtcttcacagtttggacaagaaggtccactataaagaac 7556
Db 7143 tagcccgagataagggttggatgtgtgtcttcacagtttggacaagaaggtccactataaagaac 7202
Qy 7557 gtggaccctccaaagcgaaggcgaanaaacgcttcatatagaggcgatgtgcccactacgtga 7616
Db 7203 gtggaccctccaaagcgaaggcgaanaaacgcttcatatagaggcgatgtgcccactacgtga 7262
Qy 7617 ccataccaccttaataaatttcttcttggtgtcagagtgcgttaaagcactaaatcgaacccct 7676
Db 7263 ccataccaccttaataaatttcttcttggtgtcagagtgcgttaaagcactaaatcgaacccct 7322
Qy 7677 aaagggagcccccgatcttaagctttgaagtcgagggaagcggcggaacgctgycgagaagaagaa 7736
Db 7323 aaagggagcccccgatcttaagctttgaagtcgagggaagcggcggaacgctgycgagaagaagaa 7382
Qy 7737 gggaagaagaagcgaagaagcggcggtcagagcgctgagcgctgttaacgctgcacagctgcgc 7796
Db 7383 gggaagaagaagcgaagaagcggcggtcagagcgctgagcgctgttaacgctgcacagctgcgc 7442
Qy 7797 gtaaccacacacacccgcgcgcgttaatgctgcgcgtacagaagcgct 7842
Db 7443 gtaaccacacacacccgcgcgcgttaatgctgcgcgtacagaagcgct 7488

RESULT 3
US-09-470-618-14
: Sequence 14, Application US/09470618
: Patent No. 6200560
: GENERAL INFORMATION:
: APPLICANT: Couto, Linda B.
: APPLICANT: Colosi, Peter C.
: TITLE OF INVENTION: Adeno-associated Vectors for Expression of Factor VIII
: FILE REFERENCE: by Target Cells
: FILE REFERENCE: Avigen-04082
: CURRENT APPLICATION NUMBER: US/09/470,618
: CURRENT FILING DATE: 1999-12-22
: EARLIER APPLICATION NUMBER: 09/364,862
: EARLIER FILING DATE: 1999-07-30
: EARLIER APPLICATION NUMBER: 60/125,974
: EARLIER FILING DATE: 1999-03-24
: EARLIER APPLICATION NUMBER: 60/104,994
: EARLIER FILING DATE: 1998-10-20
: NUMBER OF SEQ ID NOS: 15
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 14
: LENGTH: 4999
: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-470-618-14

```

	Best Local Similarity	93.7%	Pred. No. 0;	
	Matches 4/24;	Conservative	0; Mismatches 243;	Indels 72; Gaps 5;
QY	15	ctgcgcgctctgcctcgtctcaactgaagcccgaggacaaagtgcgcgcagcccgagctt	74	
Db	16	ctgcgcgctctgcctcgtctcaactgaagcccgaggacaaagcccgagcttgcgcgacctt	75	
QY	75	gcccgagcgccctcaagtgagcgaagcgaagcgcgcagagagaggagtgycacaatccatcact	134	
Db	76	gtctgcgccgcgtccatcagtgagcgaagcgaagcgaagagaggagtgycacaatccatcact	135	
QY	135	aggggtcttccatgaatctcttcttaagtaaaagtaactgaacctttaccccgttgctcgg	194	
Db	136	aggggtcttccgcgcgcgcgcgcgcgtgtrt-ggcgcgggggttaaacctgagaaagtgaatgctcg	194	
QY	195	caaaagcctctgctctgtgcgaagtgcttctgtcagcgaaccccccacatctgtcgcttggcc	254	
Db	195	tactgtgcctccgctcttctcccgagagtgggggagaaacgctatataagtgcaatgctgcgc	254	
QY	255	ataaggccatcagcgcagtgagatctcagctgtggtgttttcgaagagaaagcaaaaagcctct	314	
Db	255	gtgaacgctcttcttcgcgaacggttttgccgcgcgcgcgcgcagtgcaagtgccagggaa	314	
QY	315	ccaaaccagcgcttggaatgttctccaaaccaaagtccgaagcagtggtgttttcgaagaggaagc	374	
Db	315	tgtgtcttaaaatcacatgcgtccaggggaatgt-----tgtgtcttaaatcacatcact	368	
QY	375	aaaaagcctctccaaaccgaagcgcttgagctcggagagc--tgcgaacacatgccaatatagagc	432	
Db	369	gacactgacacaccaaacttttctcttcttcccccagagtatcgatccacaatgcaaatlaagc	428	
QY	433	tctccacactgtctcttctctgtgccttttgcgacttctgacttcttaagtgccacaagaagatact	492	
Db	429	tctccacactgtctcttctctgtgccttttgcgacttctgacttcttaagtgccacaagaagatact	488	
QY	493	acctgtgtgcagtggaactgtcattgaggaactatattgcaaaatgtatctcgtgtgagctgcctg	552	
Db	489	aacctgtgtgcagtggaactgtcattgaggaactatattgcaaaatgtatctcgtgtgagctgcctg	548	
QY	553	tggaaagcagaagttccctcctcctagatgagtcgcaaaaatcttttccatcacaacacttcagtcgt	612	
Db	549	tggaaagcagaagttccctcctcctcctagatgagtcgcaaaaatcttttccatcacaacacttcagtcgt	608	
QY	613	acaaaaaagacactctgttctgtgaagaattcacacggtttcacaccttttcaacatctgcctcaagccaagc	672	
Db	609	acaaaaaagacactctgttctgtgaagaattcacacggtttcacaccttttcaacatctgcctcaagccaagc	668	
QY	673	caaccttgagtggtctctgactaagtccttaacatccoaagcttgaggtttatgatacagtggtca	732	
Db	669	caaccttgagtggtctctgactaagtccttaacatccoaagcttgaggtttatgatacagtggtca	728	
QY	733	ttacacttaaaaaaataatgcttcccatcctgttcagcttctcaatgtctgtgtgtgtactact	792	
Db	729	ttacacttaaaaaaataatgcttcccatcctgttcagcttctcaatgtctgtgtgtgtactact	788	
QY	793	ggaagaactctcgaagggagcctgaatatatgatcatcagaccgaatccaagggagaaagaaagatg	852	
Db	789	ggaaagactctcgaagggagcctgaatatatgatcatcagaccgaatccaagggagaaagaaagatg	848	
QY	853	ataaagctcttccctgtgtgtgaaagcacaatatatgtctgcgcagctctctgnaaagagaatgtct	912	
Db	849	ataaagctcttccctgtgtgtgaaagcacaatatatgtctgcgcagctctctgnaaagagaatgtct	908	
QY	913	caatgagcctcgaacccacatgctgccttaccctaccatcaatacttcttcaatgttgagactgtta	972	
Db	909	caatgagcctcgaacccacatgctgccttaccctaccatcaatacttcttcaatgttgagactgtta	968	
QY	973	aagaacttaattccagcgctcatctggaagccctactagatgtttaagaagaggagctcgtgcca	1032	
Db	969	aagaacttaattccagcgctcatctggaagccctactagatgtttaagaagaggagctcgtgcca	1028	
QY	1033	aggaagaagacacagaccttgcaacaatttactactcttctgtctgatatattgatgaaaggga	1092	



Db 1029 aggaaaagacacagaccttgcacaaattatactcttctgctgatttgaagga 1088  
QY 1093 aaagtgtgacctcagaacaaagaactccttgatgcagatagagatgtgcatctgcc 1152  
Db 1089 aaagtgtgacctcagaacaaagaactccttgatgcagatagagatgtgcatctgcc 1148  
QY 1153 gggccttgacctaaatgacacagatcaatgtgtatgttaaacaggtctctgcaggtctga 1212  
Db 1149 gggccttgacctaaatgacacagatcaatgtgtatgttaaacaggtctctgcaggtctga 1208  
QY 1213 ttgatatgcacagagaaatcagatctatgttgatgttgaaatgggacccactctgaag 1272  
Db 1209 ttgatatgcacagagaaatcagatctatgttgatgttgaaatgggacccactctgaag 1268  
QY 1273 tgcactcacaatatctctgcgaaggtcacacatctctgttgaggaacacacgcgcgctct 1332  
Db 1269 tgcactcacaatatctctgcgaaggtcacacatctctgttgaggaacacacgcgcgctct 1328  
QY 1333 ttggaatctgcacaaataacttctctactgtctcaaacactcttgatggaccttgaagat 1392  
Db 1329 ttggaatctgcacaaataacttctctactgtctcaaacactcttgatggaccttgaagat 1388  
QY 1393 ttctactggtttgtcatatctctctctcccaacacatgataggcatggaaagtattgtccaag 1452  
Db 1389 ttctactggtttgtcatatctctctctcccaacacatgataggcatggaaagtattgtccaag 1448  
QY 1453 taagacagctgtccagaggaaccccacacacagatgaaaaataatgaagaagcggagact 1512  
Db 1449 taagacagctgtccagaggaaccccacacacagatgaaaaataatgaagaagcggagact 1508  
QY 1513 atgatatgatacttaactgaatctctgaaatgatagtgtgcaaggtctgatatgaacactctc 1572  
Db 1509 atgatatgatacttaactgaatctctgaaatgatagtgtgcaaggtctgatatgaacactctc 1568  
QY 1573 ctctcttaaccacaaattgcctcgaagtctgcagaagaagcatcccttaaacacttgggtacattaa 1632  
Db 1569 ctctcttaaccacaaattgcctcgaagtctgcagaagaagcatcccttaaacacttgggtacattaa 1628  
QY 1633 ttgctgtctgaagaaggagactctgagactatgtctcccttagtccctgcgccccgagtacaga 1692  
Db 1629 ttgctgtctgaagaaggagactctgagactatgtctcccttagtccctgcgccccgagtacaga 1688  
QY 1693 gtataaaagtcacaatatcttgaaacaaatggcctcagcgatgtgtctggaagtacaaaag 1752  
Db 1689 gtataaaagtcacaatatcttgaaacaaatggcctcagcgatgtgtctggaagtacaaaag 1748  
QY 1753 ttccattatgtgcataccagatgaaacacttaagactcgtgaagcttatcttcagctgaat 1812  
Db 1749 ttccattatgtgcataccagatgaaacacttaagactcgtgaagcttatcttcagctgaat 1808  
QY 1813 caggaaatcttggaaccttacttatactgggaagtctgagacacacatgttgcataatata 1872  
Db 1809 caggaaatcttggaaccttacttatactgggaagtctgagacacacatgttgcataatata 1868  
QY 1873 agaatcaagcagacagacacataatacattacacctcaccggaatacattgtctcgtctct 1932  
Db 1869 agaatcaagcagacagacacataatacattacacctcaccggaatacattgtctcgtctct 1928  
QY 1933 ttgatatgaagagattacccaagaagtgttaaaaacatttgaagaatttccaattctgcag 1992  
Db 1929 ttgatatgaagagattacccaagaagtgttaaaaacatttgaagaatttccaattctgcag 1988  
QY 1993 gagaaatatataataataatgacagatgactgttagaagatgggccaactaaatcacatc 2052  
Db 1989 gagaaatatataataataatgacagatgactgttagaagatgggccaactaaatcacatc 2048  
QY 2053 ctggtgtcctgaacccgctataactctagtctcgttaataatgagagagatctcagcttaag 2112  
Db 2049 ctggtgtcctgaacccgctataactctagtctcgttaataatgagagagatctcagcttaag 2108  
QY 2113 gaacctatggcctctctctactctgctacaaagaattcgttagataaagagagaacccaga 2172  
Db 2109 gaacctatggcctctctctactctgctacaaagaattcgttagataaagagagaacccaga 2168

QY 2173 taatgtcagacagaagaatgtcatccctgtttctctgtatatttgatgagaaaccgaagctggt 2232  
Db 2169 taatgtcagacagaagaatgtcatccctgtttctctgtatatttgatgagaaaccgaagctggt 2228  
QY 2233 acctacagaagaatatacaacgcttctctcccaatccagctcggagatgcagcttgaagatc 2292  
Db 2229 acctacagaagaatatacaacgcttctctcccaatccagctcggagatgcagcttgaagatc 2288  
QY 2293 cagagttccaaagcctccaaatcatgcagatgcagaatcaaggtatgtttttgatatgttgc 2352  
Db 2289 cagagttccaaagcctccaaatcatgcagatgcagaatcaaggtatgtttttgatatgttgc 2348  
QY 2353 agttgtcagttgtttgtgatagggtgtgcatactgtgtatacttcaagcatcttgagacacaga 2412  
Db 2349 agttgtcagttgtttgtgatagggtgtgcatactgtgtatacttcaagcatcttgagacacaga 2408  
QY 2413 ctgacttctctctgtctctctctctgtgatatactctaaacacaaatggtctataag 2472  
Db 2409 ctgacttctctctgtctctctctctgtgatatactctaaacacaaatggtctataag 2468  
QY 2473 acacactaacccatctccatctcaggaagaactgtcttcatgtcgttgaaaaaccag 2532  
Db 2469 acacactaacccatctccatctcaggaagaactgtcttcatgtcgttgaaaaaccag 2528  
QY 2533 gtctatgatatctgggtgtgcacaaactcagaactcttcggaacagaggtcagtcgcgcttac 2592  
Db 2529 gtctatgatatctgggtgtgcacaaactcagaactcttcggaacagaggtcagtcgcgcttac 2588  
QY 2593 tgaaggttctagctgtgtgaacaaacactgtgtgatatatacgaagagacgtatagaata 2652  
Db 2589 tgaaggttctagctgtgtgaacaaacactgtgtgatatatacgaagagacgtatagaata 2648  
QY 2653 ttccagcatctgtctgtgataaacaacatgtccatgtgaacaaagagctctccagatatt 2712  
Db 2649 ttccagcatctgtctgtgataaacaacatgtccatgtgaacaaagagctctccagatatt 2707  
QY 2713 caagacaccctagcacatagcaaaaagcaatttaatgtccaccaccagctcttgaacgccc 2772  
Db 2708 -----ccaccagctcttgaacgccc 2726  
QY 2773 atcaacgggaaataactcgttactactcttcagtcagatccaagagaaattgacataatg 2832  
Db 2727 atcaacgggaaataactcgttactactcttcagtcagatccaagagaaattgacataatg 2786  
QY 2833 ataccatatcagttgaatgaaatgaaagagaaatttggacattatgataagatgaaatc 2892  
Db 2787 ataccatatcagttgaatgaaatgaaagagaaatttggacattatgataagatgaaatc 2846  
QY 2893 agagcccccgcagcttccaagaagaacacagacataatttatgtctgcagtggagaagc 2952  
Db 2847 agagcccccgcagcttccaagaagaacacagacataatttatgtctgcagtggagaagc 2906  
QY 2953 tctgtgatatgaggtatgtagtgcctcccaacatgttctaagaacaaagggcttaagatgtgca 3012  
Db 2907 tctgtgatatgaggtatgtagtgcctcccaacatgttctaagaacaaagggcttaagatgtgca 2966  
QY 3013 gttgcccccagttcaagaagaattgtttccagggaaatttacgtaggcctcttactcagc 3072  
Db 2967 gttgcccccagttcaagaagaattgtttccagggaaatttacgtaggcctcttactcagc 3026  
QY 3073 ccttataccgttgaagaataatgaaacatttggagactcctgtgggcatataataagagcag 3132  
Db 3027 ccttataccgttgaagaataatgaaacatttggagactcctgtgggcatataataagagcag 3086  
QY 3133 aagttgaagaataataatcaatgttaacttccaagaatcagggcctctcgtccctatctctct 3192  
Db 3087 aagttgaagaataataatcaatgttaacttccaagaatcagggcctctcgtccctatctctct 3146  
QY 3103 attcagccttattcttatgagagaatcagagcgaagcgaagacccatagaaaaaact 3252  
Db 3147 attcagccttattcttatgagagaatcagagcgaagcgaagacccatagaaaaaact 3206



QY	3253	tttcaagccctaagaacccaacacactttgggaagtgcacaatcatatgagacca	3312
Db	3207	ttgtcaagccctaagaacccaacacttcttggaaagtgcacaatcatatgagacca	3266
QY	3313	ctaaagatgagtttgactgcgaagccctgggcttattcttcctgatgttgcacctggaaaag	3372
Db	3267	ctaaagatgagtttgactgcgaagccctgggcttattcttcctgatgttgcacctggaaaag	3326
QY	3373	atgtgcactcagcgctgatttgcagcccttcctgtgtcgcacactaacaactgaaacctg	3432
Db	3327	atgtgcactcagcgctgatttgcagcccttcctgtgtcgcacactaacaacactgaaacctg	3386
QY	3433	ctcatggagagcaagctgacagctacaggaatttgctctgttttccacactttgataga	3492
Db	3387	ctcatggagagcaagctgacagctacaggaatttgctctgttttccacactttgataga	3446
QY	3493	ccaagaagctgttacttccacgtaaatatgtgaaagaacctgcagcgctccctgcgaatacc	3552
Db	3447	ccaagaagctgttacttccacgtaaatatgtgaaagaacctgcagcgctccctgcgaatacc	3506
QY	3553	agatggaagatcccaacttttaagaagaaatatactgccttccatgcaatcaatggtacataa	3612
Db	3507	agatggaagatcccaacttttaagaagaaatatactgccttccatgcaatcaatggtacataa	3566
QY	3613	tgtgatacacctacgctgctctgtatagtgtccaggaatcaaaagatctgatgtatctgtca	3672
Db	3567	tgtgatacacctacgctgctctgtatagtgtccaggaatcaaaagatctgatgtatctgtca	3626
QY	3673	gcatgtggagaaatgaaacaatccattcttcatcttaagtgagaaatgttctacctgtac	3732
Db	3627	gcatgtggagaaatgaaacaatccattcttcatcttaagtgagaaatgttctacctgtac	3686
QY	3733	gaaaaaaagagagatataaaatgycacgtgtacaaatctccatccacaagttttttgagaacg	3792
Db	3687	gaaaaaaagagagatataaaatgycacgtgtacaaatctccatccacaagttttttgagaacg	3746
QY	3793	tgtgaaatgttacaatccaagctgcgaattgtgcggtgtgaaatgtcctattgtgcgagacac	3852
Db	3747	tgtgaaatgttacaatccaagctgcgaattgtgcggtgtgaaatgtcctattgtgcgagacac	3806
QY	3853	tacatgtcgggaatgagacaactttttctgtgtgtacagcaataagtgtaagctccctcg	3912
Db	3807	tacatgtcgggaatgagacaactttttctgtgtgtgtacagcaataagtgtaagctccctcg	3866
QY	3913	gaatgtcctcgcgacacatatagagattttcagaattacaagcttccagacaatatatgagacgt	3972
Db	3867	gaatgtcctcgcgacacatatagagattttcagaattacaagcttccagacaatatatgagacgt	3926
QY	3973	gggcccacaagctgtgccagactcattatctcggaatcaatcaatgtcctgtgcaccacaag	4032
Db	3927	gggcccacaagctgtgccagactcattatctcggaatcaatcaatgtcctgtgcaccacaag	3986
QY	4033	agcccttttttggataaagtgtagctgtgtgcacccaatgatatcttaacgtgcataaga	4092
Db	3987	agcccttttttggataaagtgtagctgtgtgcacccaatgatatcttaacgtgcataaga	4046
QY	4093	cccaagggtgcocgtcgaagaagtctccagacctcaaatctcctaagtttaactgaatgata	4152
Db	4047	cccaagggtgcocgtcgaagaagtctccagacctcaaatctcctaagtttaactgaatgata	4106
QY	4153	gtcttgaatgggaagaagtgtgcagactatcgcsgaataatccaactgaaaccttaatgtct	4212
Db	4107	gtcttgaatgggaagaagtgtgcagactatcgcsgaataatccaactgaaaccttaatgtct	4166
QY	4213	tctttggcaatgttgatatacttggtgataaacaacaatatattttaacctccaattatgt	4272
Db	4167	tctttggcaatgttgatatacttggtgataaacaacaatatattttaacctccaattatgt	4226
QY	4273	ctcgatcacatccgtttgcaccaaccaactatataagcatctgcagcaactcttcgaatgagt	4332
Db	4227	ctcgatcacatccgtttgcaccaaccaactatataagcatctgcagcaactcttcgaatgagt	4286
QY	4333	tgtatggcctgtgtatttaaatagtgttcgacatgtccattgtggaatggaggtlaaagcaatat	4392

Db	4287	tgatggtcgtgatttaataagcttgcagcagctatcttggagatggagataaagcaat	4346
Qy	4393	cgagatgcacagattactgtcttcaccccttaccatatttttgcacccgtgtcctt	4452
Db	4347	cgagatgcacagattactgtcttcaccccttaccatatttttgcacccgtgtcctt	4406
Qy	4453	caaaagctgcattcactccaaggagagtaatgctcgtgagacctcagtgtaatac	4512
Db	4407	caaaagctgcattcactccaaggagagtaatgctcgtgagacctcagtgtaatac	4466
Qy	4513	caaaagatgtgctgcaagtgtgacttccagaagaacaatgaaagtacagagtaactc	4572
Db	4467	caaaagatgtgctgcaagtgtgacttccagaagaacaatgaaagtacagagtaactc	4526
Qy	4573	aggagatgaatactctgtcttaccagacagatgtggaagagtttccatctccagac	4632
Db	4527	aggagatgaatactctgtcttaccagacagatgtggaagagtttccatctccagac	4586
Qy	4633	aagatgycatcatcagtgagactctcttttccagaaatggcgaagtaagttttccag	4692
Db	4587	aagatgycatcatcagtgagactctcttttccagaaatggcgaagtaagttttccag	4646
Qy	4693	atcaagactctctacacactgtgtgtgaactctctcagaccaccgttactcgtacc	4752
Db	4647	atcaagactctctacacactgtgtgtgaactctctcagaccaccgttactcgtacc	4706
Qy	4753	ttcgaattacccccacagatgttgggtgcacacagattgccttgaagatggaagttc	4812
Db	4707	ttcgaattacccccacagatgttgggtgcacacagattgccttgaagatggaagttc	4766
Qy	4813	ggcagagcacaggaactctcactgcagtcgagcagttcttctgaagggaatcgagca	4872
Db	4767	ggcagagcacaggaactctcactgcagtcgagcagttcttctgaagggaatcgagca	4805
Qy	4873	gacgaataaaaaacgacaggtgtgtgggtcgttcttgcgatccagatctagaacct	4932
Db	4806	gaaatttatcttcatctgaataagtggtgtgttcttctgtgtgcgcgcaggaacct	4865
Qy	4933	gtgagtgaggttggcacacccctctctgacgctcgtcgtcactatgagcgccggga	4992
Db	4866	gtgagtgaggttggcacacccctctctgacgctcgtcgtcactatgagcgccggga	4925
Qy	4993	aagcccgagcgctcgagcacttggatcgccggccctcagtgagcagcgagcgag	5051
Db	4926	aagtcgcgcgcagcgccgcttggccggcgccctcagtgagcagcgagcgag	4984
RESULT 4			
US-09-364-862-14			
: Sequence 14, Application US/09364862			
: Patent No. 6221349			
GENERAL INFORMATION:			
APPLICANT: Couto, Linda B.			
APPLICANT: Colosi, Peter C.			
TITLE OF INVENTION: ADENO-ASSOCIATED VECTORS FOR EXPRESSION OF FACTOR VIII			
TITLE OF INVENTION: BY TARGET			
TITLE OF INVENTION: CELLS			
FILE REFERENCE: AVIGER-03743			
CURRENT APPLICATION NUMBER: US/09/364,862			
CURRENT FILING DATE: 1999-07-30			
EARLIER APPLICATION NUMBER: 60/125,974			
EARLIER FILING DATE: 1999-03-24			
EARLIER APPLICATION NUMBER: 60/104,994			
EARLIER FILING DATE: 1998-10-20			
NUMBER OF SEQ ID NOS: 14			
SOFTWARE: PatentIn Ver. 2.0			
SEQ ID NO 14			
LENGTH: 4999			
TYPE: DNA			
ORGANISM: Artificial Sequence			
FEATURE:			
OTHER INFORMATION: Description of Artificial Sequence: Synthetic			



US-09-364-862-14

Query Match	56.18;	Score 4456.2;	DB 4;	Length 4999;
Best Local Similarity	93.78;	Pred. No. 0;		
Matches 4724; Conservative	0;	Mismatches 243;	Indels 72;	Gaps 5;

QY	15	ctgagcgctcgcgtctgcacacatctgagcgagcgagcaaaagtcggccgagcgccggcctt	74
Db	16	ctgagcgctcgcgtctgcacacatctgagcgagcgagcaaaagtcggccgagcgccggcctt	75
QY	75	gccccgggcgagcctcgaatgtagcgagcgagcgagcgagagtagtggtgcacacccatacct	134
Db	76	gtctgcgccgagcctcgaatgtagcgagcgagcgagagtagtggtgcacacccatacct	135
QY	135	agggagttctctcgaatctctcttctcaagtaaacgatacgaaccttaccgcgtctcgg	194
Db	136	agggagttctctcgaatctctcttctcaagtaaacgatacgaaccttaccgcgtctcgg	194
QY	195	caaaagcctctgctcgtctcgaagtgcttctgcagcagcaaccccccctgcgtggccttggcc	254
Db	195	tactgcgtccgcgccttctcccgaggtgtggggagaaacgtatataaagtgcagtaatgcgc	254
QY	255	ataagcagcatcagcgcagtcgagatctcagtggtgttttgcagaagagagcaaaaagcctct	314
Db	255	gtgaaagcttctcttctcgcagaacggtgtgtgcgcccgccgcgcgcaggtaaagtgtccaggaatgt	314
QY	315	ccaccacgcgccttggaatggttccacacacacatgctcagcagtggtgttttgcagaagagc	374
Db	315	tgttctcttaaatatacacaatcgcgcgcagggaaagt-----tgttcttaaatatacacaatcct	368
QY	375	aaaaagcctctcaccacacacgccttgagactgcagagc--ctgcagacacacatgcacataagc	432
Db	369	gacactgacatacacaacttctctcttcttcacagagatcgcataccacatgcacataagc	428
QY	433	tctccacctgctctctctctctgtgccttttgatcttcgctttggtgcacacagaatact	492
Db	429	tctccacctgctctctctctctgtgccttttgatcttcgctttggtgcacacagaatact	488
QY	493	accctgggtgcagctgagaaactgcatactgagacatactgcaaaagtgcctcgcgtgcagctgt	552
Db	489	accctgggtgcagctgagaaactgcatactgagacatactgcaaaagtgcctcgcgtgcagctgt	548
QY	553	tggagcgcagaatcttctcctctcctagagtgccaaatcttccatctcaacacactcagtcggt	612
Db	549	tggagcgcagaatcttctcctcctcctagagtgccaaatcttccatctcaacacactcagtcggt	608
QY	613	acaaaagagactctgtttgtagaatctcaagcttcaacttctcaacatcgcgttaagcaagc	672
Db	609	acaaaagagactctgtttgtagaatctcaagcttcaacttctcaacatcgcgttaagcaagc	668
QY	673	caaccctgagatggtctcgtcagtgctcacaatccacagctgaggtcttaagatacagtggtca	732
Db	669	caaccctgagatggtctcgtcagtgctcacaatccacagctgaggtcttaagatacagtggtca	728
QY	733	ttaacacttaagaaatgagcttccacatccgttcagtgcttcaatgctgttgggtgatactact	792
Db	729	ttaacacttaagaaatgagcttccacatccgttcagtgcttcaatgctgttgggtgatactact	788
QY	793	ggaagactctcgagggagcctgaaatagatgatcatagacccgtaacaaagggtgaagaagaatg	852
Db	789	ggaagactctcgagggagcctgaaatagatgatcatagacccgtaacaaagggtgaagaagaatg	848
QY	853	ataaagctctccctcgtgtggaagccatacatalatgtctgcaggtccctbgaagaagaatggtc	912
Db	849	ataaagctctccctcgtgtggaagccatacatalatgtctgcaggtccctbgaagaagaatggtc	908
QY	913	caatggcctctcgaccacactgtgccttaacacacacatacttctctatgtggagcctgtgtaa	972
Db	909	caatggcctctcgaccacactgtgccttaacacacacatacttctctatgtggagcctgtgtaa	968
QY	973	aaagactgaaatcgaagcctcaatctggagccctactagatagtgtagaagaggagatctgtgcca	1032

Dh	969	aagacttgaattccagcccttaatttgagcccttaagtatgtagagaagagatctgcgcca	1028
Qy	1033	aggaanaagacacagaccttgcacaaattatactacttttgcgtatattgataagaagga	1092
Dh	1029	aggaanaagacacagaccttgcacaaattatactacttttgcgtatattgataagaagga	1088
Qy	1093	aaagtctgcactcagaaacaaagaactccttgaatgcagaatagggatgctgcatactgc	1152
Dh	1089	aaagtctgcactcagaaacaaagaactccttgaatgcagaatagggatgctgcatactgc	1148
Qy	1153	gggctctgccttaaatatgcacacagctoaatggttatgttaaacaggtctctgcacagtcga	1212
Dh	1149	gggctctgccttaaatatgcacacagctoaatggttatgttaaacaggtctctgcacagtcga	1208
Qy	1213	ttgagtgcacacaggaataatgcagctcatgtgatgtgataatgtgaatgcacactcttcgaag	1272
Dh	1209	ttgagtgcacacaggaataatgcagctcatgtgatgtgataatgtgaatgcacactcttcgaag	1268
Qy	1273	tgcaactcaatatccctcogaaggtcacacaattctctgtgaagaaacatcgcagcgctcc	1332
Dh	1269	tgcaactcaatatccctcogaaggtcacacaattctctgtgaagaaacatcgcagcgctcc	1328
Qy	1333	tggaaatctgcgcgaataacttcttcatactgtcttaaacacactttgtatggagccttgaacgt	1392
Dh	1329	tggaaatctgcgcgaataacttcttcatactgtcttaaacacactttgtatggagccttgaacgt	1388
Qy	1393	ttctactgttttctgcatacatctctccccaacaaatgatgtgcataagaagcttatgtccaag	1452
Dh	1389	ttctactgttttctgcatacatctctccccaacaaatgatgtgcataagaagcttatgtccaag	1448
Qy	1453	tagaagagctgtccagaggaacccccaactacgaaatgaanaataatgaagaagcggagaagct	1512
Dh	1449	tagaagagctgtccagaggaacccccaactacgaaatgaanaataatgaagaagcggagaagct	1508
Qy	1513	atgtagatgatattactacgattcttcgaaatagatgtagtgcgtgcgttatgatgaacaacttc	1572
Dh	1509	atgtagatgatattactacgattcttcgaaatagatgtagtgcgtgcgttatgatgaacaacttc	1568
Qy	1573	ctcccttatcccaaatcgcctcagcttgcccaagaagatcccttaaaccttggatatactaca	1632
Dh	1569	ctcccttatcccaaatcgcctcagcttgcccaagaagatcccttaaaccttggatatactaca	1628
Qy	1633	ttgcgtcgtgaagagagagactgvggactatgctcccttaagtcctcgcgcccgatagcaga	1692
Dh	1629	ttgcgtcgtgaagagagagactgvggactatgctcccttaagtcctcgcgcccgatagcaga	1688
Qy	1693	gtttataaagatcaatattttgaacaatvgccctcagaagcttggatggagaatgaacaaaag	1752
Dh	1689	gtttataaagatcaatattttgaacaatvgccctcagaagcttggatggagaatgaacaaaag	1748
Qy	1753	tcgcattatbgcatalacacagatgaaaccttgaagctcgtttaagctatcagcagat	1812
Dh	1749	tcgcattatbgcatalacacagatgaaaccttgaagctcgtttaagctatcagcagat	1808
Qy	1813	caggaaatcttggagaccttacttaattatvgggaggttggaaacacactgttgatatattta	1872
Dh	1809	caggaaatcttggagaccttacttaattatvgggaggttggaaacacactgttgatatattta	1868
Qy	1873	agaatacagaacagacagacatatatacatctacccctcagaagctcagatgctgcctc	1932
Dh	1869	agaatacagaacagacagacatatatacatctacccctcagaagctcagatgctgcctc	1928
Qy	1933	tgatttaagagagatttcacaaaggtgttaaaaattttgaagaatttccaattctgcag	1992
Dh	1929	tgatttaagagagatttcacaaaggtgttaaaaattttgaagaatttccaattctgcag	1988
Qy	1993	gagaataattccaataataatgaatggacagtgagctttaagaatgtggccaactaatcaatgc	2052
Dh	1989	gagaataattccaataataatgaatggacagtgagctttaagaatgtggccaactaatcaatgc	2048
Qy	2053	ctcgtgcctgcagcccgctattactcagtttcgttaataatggagaagatcagctcag	2112
Dh	2049	ctcgtgcctgcagcccgctattactcagtttcgttaataatggagaagatcagctcag	2108



QY	2113	gaccatttggccctctctcattctgtctacaagaatctgtatgatacaaaagagaaaccaga	2172
Db	2109	gactcatttggccctctctcattctgtctacaagaatctgtatgatacaaaagagaaaccaga	2168
QY	2173	taatgtcaagaagaagaatgatactctgtttttctgtatttgaatgaacccgaagctggt	2232
Db	2169	taatgtcaagaagaagaatgatactctgtttttctgtatttgaatgaacccgaagctggt	2228
QY	2233	accctcaagaataatatacaacgcctctctcccaatccagctgaagtgaacgttgagatc	2292
Db	2229	accctcaagaagaataatatacaacgcctctctcccaatccagctgaagtgaacgttgagatc	2288
QY	2293	cagaagttccaaagcctccacaatcatgacaagcatcaatgagctatgtttttgataagttgc	2352
Db	2289	cagaagttccaaagcctccacaatcatgacaagcatcaatgagctatgtttttgataagttgc	2348
QY	2353	agttgtcaagtttcttgatgagaagtgtgcatactggtatacttcaagcatvtgagacaaga	2412
Db	2349	agttgtcaagtttcttgatgagaagtgtgcatactggtatacttcaagcatvtgagacaaga	2408
QY	2413	ctgagctcccttcctgctctctctctccgggataaccttcaaaacaanaatgttctatgaag	2472
Db	2409	ctgagctcccttcctgctctctctctccgggataaccttcaaaacaanaatgttctatgaag	2468
QY	2473	acaacatccaccttatloccattctcagagagaactgtctctcatgtgcgaatggaaaaaccag	2532
Db	2469	acaacatccaccttatloccattctcagagagaactgtctctcatgtgtcgaatggaaaaaccag	2528
QY	2533	gtctatgattcttgggggtgcacaaactcgaactcttcogaaacagagcatgaacgcctttac	2592
Db	2529	gtctatgattcttgggggtgcacaaactcgaactcttcogaaacagagcatgaacgcctttac	2588
QY	2593	tgaaggtgttccttaatttggaaagaacaacgctggtattatctacagagacaagttataagata	2652
Db	2589	tgaaggtgttccttaatttggaaagaacaacgctggtattatctacagagacaagttataagata	2648
QY	2653	tttaagaataactgtctgtgaatgaanaaacaatgtgcaatttgaaccaagaagctcttcccaagatt	2712
Db	2649	tttaagaataactgtctgtgaatgaanaaacaatgtgcaatttgaaccaagaagctcttcccaagatt	2707
QY	2713	caagaaaccttagcactaagtcgaaagaacttaatgtccaccccaagctcttgaacgcct	2772
Db	2708	caagaaaccttagcactaagtcgaaagaacttaatgtccaccccaagctcttgaacgcct	2726
QY	2773	atcaaacgggaaataaactcgtatcaactccttcacgtcagatcaagaaggaataatcaatcatgtg	2832
Db	2727	atcaaacgggaaataaactcgtatcaactccttcacgtcagatcaagaaggaataatcaatcatgtg	2786
QY	2833	atacaataataagttgaatgaagaagaagaatttggacattatataatagaatgaataatc	2892
Db	2787	atacaataataagttgaatgaagaagaagaatttggacattatataatagaatgaataatc	2846
QY	2893	agagccccgcagccttcaaaagaanaacagacactatttattgtctgtcagvtgagaagtc	2952
Db	2847	agagccccgcagccttcaaaagaanaacagacactatttattgtctgtcagvtgagaagtc	2906
QY	2953	tcctgggtatttggatgagatgagctcccccaatgcttctaagaacacgggtctcaagvtggca	3012
Db	2907	tcctgggtatttggatgagatgagctcccccaatgcttctaagaacacgggtctcaagvtggca	2966
QY	3013	gtgccccctgaagttccaagaagaatgtgtttctccaggaatttactgaatgagctcccttaacgc	3072
Db	2967	gtgccccctgaagttccaagaagaatgtgtttctccaggaatttactgaatgagctcccttaacgc	3026
QY	3073	ccctataacgttggagaactaaatgtgaacatttggagctctctggggccatataataagagcag	3132
Db	3027	ccctataacgttggagaactaaatgtgaacatttggagctctctggggccatataataagagcag	3086
QY	3133	aagtttgaagaataatacatggtgtaactcttcaagaatccagagcctctctgtctccattctctc	3192
Db	3087	aagtttgaagaataatacatggtgtaactcttcaagaatccagagcctctctgtctccattctctc	3146

QY	3193	atctagacctatttcttatatgaggaagatcagaagcgaaaggacagaacctagaanaact	3252Z
Db	3147	attctagacctatttcttatatgaggaagatcagaagcgaaaggacagaacctagaanaaact	3206Z
QY	3253	ttgtcaagccttaattgaaacccaacctacttttggaaagtgaacaactaatatggaccca	3312.2Z
Db	3207	ttgtcaagccttaattgaaacccaacctacttttggaaagtgaacaactaatatggaccca	3266Z
QY	3313	ctaaagatgagatttggactgccaagacgtcgggtctatttctctgatatgtgactcgtgaanaag	3372Z
Db	3267	ctaaagatgagatttggactgccaagacgtcgggtctatttctctgatatgtgactcgtgaanaag	3376Z
QY	3373	atgtgactcgaagcctgtatggacctctgtgtctgtccacaactacaacatgaaacctgt	3432Z
Db	3327	atgtgactcgaagcctgtatggacctctgtgtctgtccacaactacaacatgaaacctgt	3386Z
QY	3433	ctcatgggaagacaagtgaacagttaacaggaatttctctgtttttccaccttttattgta	3492Z
Db	3387	ctcatgggaagacaagtgaacagttaacaggaatttctctgtttttccaccttttattgta	3446Z
QY	3493	ccaagaagcttggtaactctacatgaaataatggaagaagaactcgcagggctccctcgaataacc	3555Z
Db	3447	ccaagaagcttggtaactctacatgaaataatggaagaagaactcgcagggctccctcgaataacc	3506Z
QY	3553	agatggaagatcccaacttttaagaagaaatatatcgtcttcatacgaatcaaatgtgtcataaa	3612Z
Db	3507	agatggaagatcccaacttttaagaagaaatatatcgtcttcatacgaatcaaatgtgtcataaa	3566Z
QY	3613	ttgatacaactacgtgcttgaatgtgtcagaatccaagaagattcgaatgtatctgtcca	3672Z
Db	3567	ttgatacaactacgtgcttgaatgtgtcagaatccaagaagattcgaatgtatctgtcca	3626Z
QY	3673	gcatgvggacgaagaatgaaacaatccatctctatctatcttaagtgaaagtgtgttcactgtac	3732Z
Db	3627	gcatgvggacgaagaatgaaacaatccatctctatctctatcttaagtgaaagtgtgttcactgtac	3686Z
QY	3733	gaaaaaaagagagatataaaatgycactgtacaaactctataccaagtggttttggacag	3792Z
Db	3687	gaaaaaaagagagatataaaatgycactgtacaaactctataccaagtggttttggacag	3746Z
QY	3793	ttggaatgtttaccatccccaagaagcttggaaatttgggggttggaaatgcttatitggcggacac	3852Z
Db	3747	ttggaatgtttaccatccccaagaagcttggaaatttgggggttggaaatgcttatitggcggacac	3806Z
QY	3853	tacatgctcgggaatagacacaacttttctgtgtgtacagcaataaagtgtcagaaatccccgg	3912.2Z
Db	3807	tacatgctcgggaatagacacaacttttctgtgtgtgtacagcaataaagtgtcagaaatccccgg	3866Z
QY	3913	gaatgtgctctcgaacatatagagatttccagaattacagcttccaaagcaataatgtgacagt	3972Z
Db	3867	gaatgtgctctcgaacatatagagatttccagaatttccagaatttccaaagcaataatgtgacagt	3926Z
QY	3973	gggcccccaagctgtgccagacttcatattccggatacacaatgcctcggagacccaag	4032Z
Db	3927	gggcccccaagctgtgccagacttcatattccggatacacaatgcctcggagacccaag	3986Z
QY	4033	agccctttctcttgatacaaggttgatctgtgttgccacaatgtttatccggcatcaaga	4092Z
Db	3987	agccctttctcttgatacaaggttgatctgtgttgccacaatgtttatccggcatcaaga	4046Z
QY	4093	ccaaggtgtcccgccgacagaagtcttccagcgtctcacatccctcagttatcatcatgtata	4152Z
Db	4047	ccaaggtgtcccgccgacagaagtcttccagcgtctcacatccctcagttatcatcatgtata	4106Z
QY	4153	gtcttgatggaagaagatgtgcagaactatccaggaattccacatggaacttaatgtgtct	4212.2Z
Db	4107	gtcttgatggaagaagatgtgcagaactatccaggaattccacatggaacttaatgtgtct	4166Z
QY	4213	tcttttgcaaatgtgattctatctcgggaataaacaataattttaacctccaattatgt	4272Z
Db	4167	tcttttgcaaatgtgattctatctcgggaataaacaataattttaacctccaattatgt	4226Z
QY	4273	ctcgatcatatccgttttgcaccaaccaactatataagcatctgcagacactcttcgaatgaggt	4332.2Z



Db	4227	ctcgatacattccgtttgcacccaattcatattagcatttcgcagcactcttcgcatacgagt	4286
OY	4333	tgtatgggcctgtgattttaataatgtttgcagcatgccattgggaaatgagataaagcaaat	4397
Db	4287	tgtatgggcctgtgattttaataatgtttgcagcagcatatgttcagatgagaaatgaaagcaaat	4346
OY	4393	cagatggccaaatttaccgtcttcacatccctacttccaaataatgtttgcacactgtctccct	4457
Db	4347	cagatggccaaatttaccgtcttcacatccctacttccaaataatgtttgcacactgtctccct	4406
OY	4453	caaaagctcgcacttcacactcccaaggagagataatgcttggagaccttcagatgtaataatc	4512
Db	4407	caaaagctcgcacttcacactcccaaggagagataatgcttggagaccttcagatgtaataatc	4466
OY	4513	caaaaagatgtgcctcgaagtgtgactttccagaagacatgatgaatgcacaagagataactatc	4577
Db	4467	caaaaagatgtgcctcgaagtgtgactttccagaagacatgatgaatgcacaagagataactatc	4526
OY	4573	aggagatgaataatctcctcttccacagcaatgataatgaaagatgtccctcactcccaagac	4637
Db	4527	aggagatgaataatctcctcttccacagcaatgataatgaaagatgtccctcactcccaagac	4586
OY	4633	aagatgycatcatcagtggaactctcttttcagaaatgagcaaaatgaaggttttcagaggaa	4697
Db	4587	aagatgycatcatcagtggaactctcttttcagaaatgagcaaaatgaaggttttcagaggaa	4646
OY	4693	atcaagactcttcaacactgttggatgaactctttaaaccacacggatcaactgcctac	4755
Db	4647	atcaagactcttcaacactgttggatgaactctttaaaccacacggatcaactgcctac	4706
OY	4753	tctgaattcaccccccaagatgtgggttgagaccagaatctgcctcagaaatgagaggtctcgggt	4812
Db	4707	tctgaattcaccccccaagatgtgggttgagaccagaatctgcctcagaaatgagaggtctcgggt	4766
OY	4813	gcggaggccaaagacctctactctgactctgaagcaggatcttctctgaaaggagatcgcgaataaaaa	4872
Db	4767	gcggaggccaaagacctctactctgactctgaagcaggatcttctctgaaaggagatcgcgaataaaa	4805
OY	4873	gacagaaataaaagccagcagggtgtgtgggtctgttcttcggatctccagaatcagaacccta	4937
Db	4806	gaaatttatcttcaatttgcaataagtggtgtgttttttgggtggcgccgaaagaaacccta	4866
OY	4933	gtgatggagatgtgcacactccctctctgcgcgtctgcctgcgtcaactgaaggccgcgggca	4997
Db	4866	gtgatggagatgtgcacactccctctctgcgcgtctgcctgcgtcaactgaaggccgcgggca	4926
OY	4993	aagcccgaggcgttcggaggacctgtgtgtgcgcgcggcctcaagtgaagcagaaacggcag	5051
Db	4926	aaggtgcgcgcgaacgcccggggtctgtcccgccggcctccacagttgaagcagaaacggcag	4984

US-08-415  
US-08-484-891-7  
Sequence 7, Application US/08484891  
Patent No. 5935935  
GENERAL INFORMATION:  
APPLICANT: Connolly, Sheila  
APPLICANT: Connolly, Michael  
APPLICANT: Smith, Theodore  
TITLE OF INVENTION: Aenoviral Vectors for  
TILE OF INVENTION: Treatment of Hemophilia  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSSEE: Carella, Byrne, Gillfillan,  
ADDRESSSEE: Cecchi, Stewart & Olstein  
STREET: 6 Becker Farm Road  
CITY: Roseland  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07068  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch diskette

```

1  COMPUTER:  IBM PS/2
2  OPERATING SYSTEM:  MS-DOS
3  SOFTWARE:  WordPerfect 5.1
4
5  CURRENT APPLICATION DATA:
6  APPLICATION NUMBER:  US/08/484_891
7  FILING DATE:  07-JUN-1995
8  CLASSIFICATION:  514
9
10 PRIOR APPLICATION DATA:
11 APPLICATION NUMBER:  08/218_335
12 FILING DATE:  25-MAR-1994
13 APPLICATION NUMBER:  08/074_920
14 FILING DATE:  10-JUN-1993
15
16 ATTORNEY/AGENT INFORMATION:
17 NAME:  Olstein, Elliot M.
18 REGISTRATION NUMBER:  24, 025
19 REFERENCE/DOCKET NUMBER:  271010-2733
20
21 TELECOMMUNICATION INFORMATION:
22 TELEPHONE:  201-994-1700
23
24 TELEFAX:  201-994-1744
25
26 INFORMATION FOR SEQ ID NO: 7:
27
28 SEQUENCE CHARACTERISTICS:
29
30     LENGTH:  4629 bases
31     TYPE:  nucleic acid
32     STRANDEDNESS:  single
33     TOPOLOGY:  linear
34
35     MOLECULE TYPE:  CDNA primer
36
37     FEATURE:
38
39     NAME/KEY:  Factor VIII cDNA with
40     NAME/KEY:  B domain deleted
41
42 OS-08-484-891-7

```

Query Match	55.6%	Score 4413;	DB 2;	Length 4629;
Best Local Similarity	99.8%	Pred. No. 0;		
Matches 4419; Conservative	0;	Mismatches 10;	Indels 0;	Gaps 0;

Qy	420	atggaatatagacgtccaccgtctctcttcgtgtgaccttgcattcgttttagtgc	479
Db	1	ATGGAATAGAGCTCTCCACCTGCTTCTTCTGTGCTTTTGCGATTCTGCTTAGTCC	60
Qy	480	accgaagaatatactaccgtggctgagcttggaacttcaatggagatataatgcgaagtatc	539
Db	61	ACAAGAAGATACATACCGGGTGCAGTGGAACTGTATGGGACTATATGGAAGTATCTC	120
Qy	540	ggtagactgtccctgtgagcgcgaagattccctctagatgycgaaaatccttccattcaac	599
Db	121	GGTAGACTGCTGTGGAGCGCAAGATTCTCTCAGAGTGCAGAAATCTTTTCCATTCAAC	180
Qy	600	accctcagctgtgtacaaaaaagactcgtttttagaatctaaagtttcaactttaaactc	659
Db	181	ACCTCAATCTGTATACAAAAGACTCTGTTGTATGAATTCAGGTTACCTTTTCAACATTC	240
Qy	660	gtctaaccaagccacccctgtgatgtgtctgtagagctccatccacgcagctgaagttat	719
Db	241	GCTTAGCGCAAGGCCACCCCTGGATGGGTCTGTAGTCTTACCATCCAGGCTTAGGTTAT	300
Qy	720	gatacagctgtgtcatttaacttaagaacatgctgtcccatccctgtcaattcattcgtgt	779
Db	301	GATACAGTGGTCAATTACACTTAGAACATAGGGTCCCATCTCTGCAGTTCATCTCTT	360
Qy	780	ggtgtatccctactcgtggaagactctctgtggagagcttgaaatagatgatacagacagttaaag	839
Db	361	GGTATATCTTACCTGGAAGAGCTTCTGAGGGAGCTGAATATGATGATGACGACGTAAAGG	420
Qy	840	gagagaagaatataaagctctccctgtgttgaaagcatataatgtctcgtgaagttccg	899
Db	421	GAGGAABAAGTGTATTAAGTCTTCCCTGGTGGAGCCATATATGTCTGGCAGGTCTTG	480
Qy	900	aaaagaatgtccaatgagcctctgagcccaactgtcctaactacatcatcttctcat	959
Db	481	AAAAGAATGTGCCAATGGGCTCTGACCACATGTGGCTTACTACTCATATCTTTTCAT	540
Qy	960	gttagacctgttlaaagaacttgaattcaaggcctcatgtgagccctactagatgttagagaa	1019



|||||  
Db 541 GTGACCTGGTAAAGACTGAATTGAGCCCTATTGAGCCCTACTACTATGAGAGAA 600  
Qy 1020 gggagtcctgccaaggaagaacagacacttgacaaattatactacttctgctgta 1079  
Db 601 GGGGCTCTGGCCAAAGGAAAGACACAGACCTTGCACAAATTTATACACTTTTTCGCTA 660  
Qy 1080 ttgatgaagggaagaatttgagcactcagaaacaaagaaaccccttgatgaggaatgagat 1139  
Db 661 TTTCATTAAGGAAAGTATGGCACTCAGAAACAAAGACTCTTGATGAGAGATAGGAT 720  
Qy 1140 gctgcatctgctcgggcctgagctaaatgacacagtcacttgatgtaaaagagctct 1199  
Db 721 GCTGATCTGCTGGGCGCTGGCCCTAAATGCACACAGTCAATGATTATGTAACAGCTCT 780  
Qy 1200 ctgccaagtcctgatttgatgacacaggaatcagtcctatggcagtgtgtaagtgagc 1259  
Db 781 CTGCGAGCTGTGATGTGATCCACAGGAAATCAGTCTATTGGCATGTGATGGAAATGGC 840  
Qy 1260 accactctgaagtgcactcaatattcctcgaaggtgcacacattctctgtgaggaacat 1319  
Db 841 ACCACTCTCTAAGTGCCTCAATATTCCTGAAAGTGCACATTTCTTGAGAGAACAT 900  
Qy 1320 cgcgaagcgtctctggaatctgcgcaataacttctcctactgctcaaacactctgagt 1379  
Db 901 CGCAGGCGTCTTGGAATCTGCCCCAATACCTTCTCTACTGCTCAACACTCTTGATG 960  
Qy 1380 gaacttggaagattctctactgcttctgtaatactctcccaacaatgtagtgcagaa 1439  
Db 961 GACCTTGGAGAGTTTCTACAGTTTGTGCATATCTCTCCACCAACATGTATGGCATGGAA 1020  
Qy 1440 gcttatgtcaaaatagacagcgtctcagaagaaccccaactacgaatgtaaaataatgaa 1499  
Db 1021 GCTTATCTCAAAATAGACACTGTCCAGAGAACCCCACTAGCAATGAAAAATATGAA 1080  
Qy 1500 gaagcggaaactatgatatgattctactgattctgaaatgtagtgcaggttctgat 1559  
Db 1081 GAAGCGGAACATATGATGATCTTACTGATTCGAAATGATGATGGATGGATGTTGAT 1140  
Qy 1560 gatgaacactctctctctctctctctctctctctctctctctctctctctctctctct 1619  
Db 1141 GATGACAACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1200  
Qy 1620 tgggtacattacattgctgctgtagaagaagagacttggaactatgctcccttaagtcctgc 1679  
Db 1201 TGGGTATATTACTGTGCTGCTGAAGAGAGAGACTGGAGACTATGCTCCTTATGCTCGCC 1260  
Qy 1680 cccgatgacagaagtataaaatgcaatttgaaacatgagccctcagcggatgtgtagg 1739  
Db 1261 CCGGATGACAGAGTTATAAAAGTCAATATTGAAACATGCGCCCTCAGCGGATTGGTAGG 1320  
Qy 1740 aagtaaaaaaaatccgattatggtcatgacacagatgaaaccccttaagtcctgtaggct 1799  
Db 1321 AAGTACAAAAAAATCCGATTATGCGATACACAGATGAAACCTTTAAGACTCTGAAAGCT 1380  
Qy 1800 attcagatgaatcagaagaatctggagaccttacttaatgggaaggttggaagcacactg 1859  
Db 1381 ATTACAGATGAATCAGGAACCTTGGGACCTTTACTTTATGGGGAAGTTGGAGCACACTG 1440  
Qy 1860 tgaattatatgaagaatcagaacagacacatatataacttaccctcacygaatcact 1919  
Db 1441 TTGATTATATTATTAAGATCAAGCAGACATATTAACATCTACCCCTCAGGAATCACT 1500  
Qy 1920 gatgctgctctctctctctctctctctctctctctctctctctctctctctctctctct 1979  
Db 1501 GATGCTCGCTCTTTGTATTAAGAGATTACCAAAAGGTTATAAACATTTGGAAGATTATT 1560  
Qy 1980 ccaattctgcagaggaataatcacaataataaaggacagtgactgtagaagaatggcca 2039  
Db 1561 CCAATTCTGCAGAGAGAAATATTCAATATTAATGAGCANTGACTGTAGAGAGATGGGCCA 1620  
Qy 2040 actaatcagaatccctcggtgctgacccgctactactagttctgtaatatgagaga 2099  
|||||

Db 1621 ACTAAATCAGATCCTGGTGTGCTGACCCGCTATTACTCTAGTTTCTGTTAATATGAGAGA 1680  
Qy 2100 gatctgcttcaagactcatttggcctctctctcatctgctcaaaagaatctgtagatcaa 2159  
Db 1681 GATCTAGCTTCAGAGACTATGGCCCTCTCTCATCTGCTACAAAGAAATCTGTAGATCAA 1740  
Qy 2160 agaggaacagataatgtcagaagaagaatgtcaatccctgtctctctgatttgatgag 2219  
Db 1741 AGAGAAACCAAGTAAATATGTACAGCAAGAGAAATGTCACTCTGTTTCTGTATTTGATGAG 1800  
Qy 2220 aaccgaagctgtaacctcacagaaatatacaacgcttctccccaatccagctggagt 2279  
Db 1801 AACGGAAGCTGTACCTCAAGAAATATACAAAGCTTTCTCCCAATCAGCTGAGGTG 1860  
Qy 2280 caactggaggtaccagagttccaaagctccaacataatgacagcatcaatgctatgct 2339  
Db 1861 CAGCTTGAGGATCCAGAGTTCAGAGCTCCACATCATGACAGCATCAATGAGCTATGTT 1920  
Qy 2340 ttgtaagttgagttgagttgattgttttgatgaggtgcatgactggtatcattcttaagc 2399  
Db 1921 TTGATATGTTGCAAGTTGTGATGTTGTTTGCATGAGGTGCACTAGTATCTTAAGC 1980  
Qy 2400 attgagcagacgtactcctctctctctctctctctctctctctctctctctctctct 2459  
Db 1981 ATTGAGACAGACTGACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2040  
Qy 2460 atggtctatgaagaacacaccacccattcccaattctcagagaagaactgctcactgctg 2519  
Db 2041 ATGGTCTATGAACAACACACCTACCTATTCCTCATTTGACAGAAACATGCTTCAATGTCG 2100  
Qy 2520 atgaaacccagatctatgattctctgaggtgccaacacccagactctggaagaagagc 2579  
Db 2101 ATGGAACACCCAGCTGATGATTTCTGAGGTGCCACACACTCAGACTTTGGAACAGAGGC 2160  
Qy 2580 atgacgctctactgaaggttctagttgtgacaagaacacgtgtagattatcaagagac 2639  
Db 2161 ATGACCGCTCTACTGAAGGTTTCTAGTTGACACAGAACACTGCTGATTTATGAGAGAC 2220  
Qy 2640 agtataagaatatctcagaatactgctgtagtaaaacaatccattgagccaagaagc 2699  
Db 2221 AGTTATGAAGATATTTGAGATCTTGAGTGAATAAACAATGCCATGTGAACCAAGAACG 2280  
Qy 2700 tctcccaaatccaagaacccctcagaacacagcaagaatcttaatgacacccacca 2759  
Db 2281 TTCTCCAGATTTCAAGACACCCCTAGCACAGCAAAAGCAATTTAATGCCACCCACCA 2340  
Qy 2760 gctctgaacgcacalcaacgggaataactcgtactactctctcagtcagaatcaagaaga 2819  
Db 2341 GCTTGAAACGGCATCAACGGGAATAACTCGTACTCTTCAGTCAGATCAAGAGAA 2400  
Qy 2820 attgactatgatatcaccataatgattgaaatgaagaagaagaatttgacattatgat 2879  
Db 2401 AATGACTATGATATACCATATCATGTTGAATGAAGAAGAAATTTGACATTTATGAT 2460  
Qy 2880 gaagatgaaaatcagaagcccgcaacttccaagaagaacacagacactatattatgct 2939  
Db 2461 GAGAGTGAATAATAGAGCCCCCGACCTTTCAAAAGAAAAACAGACATTTTATTTGCT 2520  
Qy 2940 gcaatgagaagctctggaattatggaatgagtagctcccccacatgctctaagaacag 2999  
Db 2521 GCAGTGGAGAGGCTCTGGGATTTATGGATGAGTAGTCCCCACATGTTCTTAAGAAACAG 2580  
Qy 3000 gctcagaagtgagctgctcctcagtcagaagaagtgcttccaagaatttaatgtagc 3059  
Db 2581 GCTCAGAGTGAGCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2640  
Qy 3060 tctcttaaccagccctataacgctggaagaactaaatgaacatctggagactcctggggcca 3119  
Db 2641 TCTTTACACAGCCCTTATACGTTGAGAACTTAATGAACATTTGGGACTCCTGGGGCCA 2700  
Qy 3120 tatataagacagaagtgtgaagataatcatatgtagtaacttccaagaatcaagcctctgct 3179  
Db 2701 TATATGAAGACAGAAATTAAGATATATCATGTTCACTTTCAGAAATCAGGCTCTCTGCT 2760  
|||||



QY	3180	ccctatccctctcatcttcctagcctctattcttctatgaggaagtcagaggaagcaagga	3239
Db	2761	ccctattctcttcttatttcttaccctttattttcttatgagaaagtccagagcagacagaa	2820
QY	3240	ccttagaaaaaactctgcaagcctaaatgaaaccaaacttactctcttgaaagtgcacat	3299
Db	2821	cctagaaaaaaactttgtcagagccctaatgaaaccaaacttacttttggaaagtcaacat	2880
QY	3300	catatgcccceccctaaagtctggttctgactgcgaagccctgggcttattctctgagt	3359
Db	2881	caatattggcaccctctaaagatgattgttgactgcgaagccctgggcttattctctgagt	2940
QY	3360	gaacctgaaaaaagtctgcaactcaagcctgatactgacccctctctgtctgcacaactac	3419
Db	2941	gaacctgaaaaaagatgctgactcagccctgatactgacccctctctgtctgcacactaac	3000
QY	3420	acactgaaacctctcatgaggagaagaatgacagctacaggaattctgctctgttctaac	3479
Db	3001	acactgaacacctctctcatgaggagaagaatgacagctacaggaattctgctctgttctaac	3060
QY	3480	acctttgatgagacccaagaactgtaacttcaactgcaaaaataatgagaagaagaactgcaagct	3539
Db	3061	acctttgatgagacccaagaactgtaacttcaactgcaaaaataatgagaagaagaactgcaagct	3120
QY	3540	ccctgcaataatccagaatgagagaatcccaacttctaaagagaatlaatcgcttccatgcaatc	3599
Db	3121	ccctgcaaatatccagatgagagaatcccaacttctaaagagaatlaatcgcttccatgcaatc	3180
QY	3600	aatgctacataatgtgataacatacctacccctggcttgaatgctcaagatcaagaattcga	3659
Db	3181	aatgctactactaatggtatgatacactacccctggcttgaatgctcaagatcaagaattcga	3240
QY	3660	tggatctgtctgcagatgaggcaagaatgaaacatctatctatctatctcaagcggacat	3719
Db	3241	tggatctgtctgcagatgaggcaagaatgaaacatctatctatctatctcaagcggacat	3300
QY	3720	gtgttcactgtacgaaaaaaaaggagtaataaatgagcactgatacactctatccagct	3779
Db	3301	gtgttcactgtacgaaaaaaaaggagtaataataatgagcactgatacactctatccagct	3360
QY	3780	gtttttgagacagtgtgaaatgcttaccatctcaagaactggaatttggcgggtggaatgcctt	3839
Db	3361	gtttttgagacacagtgtgaaatgcttaccatctcaagaactggaatttggcgggtggaatgcctt	3420
QY	3840	atttgcgagcatctatactgtctggatgtagagcacacttctctgtgtcaagcaataatgct	3899
Db	3421	atttgcgagcagcatctatactgtctggatgtagagcacacttctctgtgtcaagcaataatgct	3480
QY	3900	cagagacccccctgggaatggtctctctgagcaaatgagaatttctgaatttaagccttcaagga	3959
Db	3481	cagagacccccctgggaatggtctctctgagcaaatgagaatttctgaatttcaagccttcaagga	3540
QY	3960	caataatgagcagtgaggcccccaaaagcttgcgcagactcatatctccggatacatcaatgcgc	4019
Db	3541	caataatgagcagtgaggcccccaaaagcttgcgcagactcatatctccggatacatcaatgcgc	3600
QY	4020	tggagaccaaaggaccccttctcttgatacaaggttgatctgttgcaccaaataatct	4079
Db	3601	tggagaccaaaggaccccttctcttgatacaaggttgatctgttgcaccaaataatct	3660
QY	4080	cagcgagataaaggccccaggttgccccctcaagaatgttccagagccctctacatcttcaagtt	4139
Db	3661	cagcgagataaaggccccaggttgccccctcaagaatgttccagagccctctacatcttcaagtt	3720
QY	4140	atcatcatgataagctcttgaatgggaagaagtgtgagaacttaccgaaggaatcttcaactgga	4199
Db	3721	atcatcatgataagctcttgaatgggaagaagaatgtgagaacttaccgaaggaatcttcaactgga	3780
QY	4200	accttaaatggtctctcttggcaatgttgatctcatctcggataaaacaacaataatcttaac	4259
Db	3781	accttaaatggtctctcttggcaatgttgatctcatctcggataaaacaacaataatcttaac	3840

QY	4260	ctcccaatattgctcgatcacatccgttgcacccaactcattataagcttcgcagact	4319
Db	3841	cctccaatatttgctccgatactacatccgttttcacccaactcattataagcttcgcagact	3900
QY	4320	cttcgcatactgagttgatgggactgctgtaattaaatagttgcacgaatgcacatgggaatgag	4379
Db	3901	cttcccaatgaggttgatgggactgctgtaattaaatagttgcacgaatgcacatgggaatgag	3960
QY	4380	agtaagaataataagaatgacagattactgcttcacccacttaccataatagttgttgc	4439
Db	3961	actaaagaataatagatgcacagattactgcttcacccacttaccataatagttgttgc	4020
QY	4440	acctgttccttcaaaaagctgcgacttcacccctccaaagggaagagataatgcttcgagacct	4499
Db	4021	acctgttccttcaaaaagctgcgacttcacccctccaaagggaagagataatgcttcgagacct	4080
QY	4500	caggtgaataatccaaaaagatggtgctgcaagtgcagtgacttcacgaagaacaatgaagtaca	4559
Db	4081	caggtgaataatccaaaaagatggtgctgcaagtgcagtgacttcacgaagaacaatgaagtaca	4140
QY	4560	ggaagtaactatccagggaagataaatctctgcttaccagatgtaagtgaagagttctc	4619
Db	4141	ggagtaactatctcagggaagataaatctctgcttaccagatgtaagtgaagagttctc	4200
QY	4620	atctccagcagtcacgaatggtccatcagttgactctcttttccagaatgycaaaagtaaag	4679
Db	4201	atctccacacacatgcaagatggtccatcagttgactctcttttccagaatgycaaaagtaaag	4260
QY	4680	gtttttcagggaataatcaagactctcttcacacactgtgtgaactctctgacccaccgtta	4739
Db	4261	gtttttcagggaataatcaagactctcttcacacactgtgtgaactctctgacccaccgtta	4320
QY	4740	ctgactcgtcactccttcgaattcacaccocagagtgtggtgcacagatgtgccttgcagatg	4799
Db	4321	ctgactcgtcactccttcgaattcacaccocagagtgtggtgcacagatgtgccttgcagatg	4380
QY	4800	gaggttctgggctgctggaagcacaaggaactctactgactgcagcgattc	4848
Db	4381	gaggttctgggctgctggaagcacaaggaactctactgactgcagcgattc	4429
RESULT 6			
US-08-683-839B-2			
Sequence 2, Application US/08683839B			
Patent No. 5/744326			
GENERAL INFORMATION:			
APPLICANT: Ili, Charles . R. et al.			
TITLE OF INVENTION: Use of Viral Cis-Acting Post-Transcriptional			
TITLE OF INVENTION: Regulatory Sequences to Increase Expression of			
TITLE OF INVENTION: Intronsless Genes Containing Near-Consensus Splice Sites			
NUMBER OF SEQUENCES: 18			
CORRESPONDENCE ADDRESS:			
ADDRESSEE: LAHIVE & COCKFIELD			
STREET: 60 State Street, suite 510			
CITY: Boston			
STATE: Massachusetts			
COUNTRY: USA			
ZIP: 02109-1875			
COMPUTER READABLE FORM:			
MEDIUM TYPE: Floppy disk			
COMPUTER: IBM PC compatible			
OPERATING SYSTEM: PC-DOS/MS-DOS			
SOFTWARE: Patentln Release #1.0, Version #1.25			
CURRENT APPLICATION DATA:			
APPLICATION NUMBER: US/08/683,839B			
FILING DATE: 11-MARCH-1996			
CLASSIFICATION: 435			
PRIOR APPLICATION DATA:			
APPLICATION NUMBER:			
FILING DATE:			
ATTORNEY/AGENT INFORMATION:			
NAME: Remillard, Jane E.			
REGISTRATION NUMBER: 38,872			











QY 4495 gaccctcagtgatgaataaccacaaagagtgctcgaagtggaactccagaagaacaaatgaag 4554  
|||||  
Db 7040 GACCTCAGGTGAATATATCCAAAAGAGTGGCTCGAGTGGAGCTTCCAAAGACATGAAAG 7099  
QY 4555 tcacagagtaactactcagaaggaataactctcgtctaccagaatglatgtgaagagt 4614  
|||||  
Db 7100 TCACAGAGTAACACTCAAGGAGTAATAATCTGCTTACACAGCATGTATGTGAAGAGT 7159  
QY 4615 tccctcattccagagtcagaatgagcatcagtggaactctctttttccagaatgagcaag 4674  
|||||  
Db 7160 TCCCATCTCCAGACAGCAAGATGGCCATCATGAGACTCTCTTTTCAAAATGGCAAG 7219  
QY 4675 taagagtttccagaggaatcaagaactccttcacacactgtgtgaactctcctagaaccac 4734  
|||||  
Db 7220 TAAAGGTTTTTCAGGGAATCAACACTCTTCACACCTGTGGTAATCTCTAGACCCAC 7279  
QY 4735 cgttactgaactcgtactcctcgaatlcacccccagaagttgtgtgcacagaattgacctga 4794  
|||||  
Db 7280 CGTTACTGACTCGCTACTCTTGAATTCACCCCGAGAGTTGGGTGCACACAGATTGCCCTGA 7339  
QY 4795 ggtatggaggttctcgtggtctgcgaaggaagaactctactgaactcgaagcagttc 4848  
|||||  
Db 7340 GGATGAGGTTCTGGGCTGGAGGACAGAGACCTTACTGAGGGGTGGCCACTGC 7393

RESULT 7  
US-08-717-294-41  
; Sequence 41, Application US/08717294  
; Patent No. 611418  
; GENERAL INFORMATION:  
; APPLICANT: SEED, BRIAN  
; APPLICANT: HAAS, JUREN  
; TITLE OF INVENTION: HIGH LEVEL EXPRESSION OF  
; TITLE OF INVENTION: PROTEINS  
; NUMBER OF SEQUENCES: 110  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Clark & Elbling LLP  
; STREET: 176 Federal Street  
; City: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02110  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/717,294  
; FILING DATE: 20-SEP-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Elbling, Karen L.  
; REGISTRATION NUMBER: 35,238  
; REFERENCE/DOCKET NUMBER: 00786/345001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-428-0200  
; TELEFAX: 617-428-7045  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 41:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4670 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; US-08-717-294-41

Query Match 54.9%; Score 4357.6; DB 3; Length 4670;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 4363; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 477 gccaccagaataactacccctgggtgcagtggaactgcatggaactatgcaaatgcat 536  
|||||  
Db 84 GCCACCAAGAAATACCTACCTGGGTGACAGTGAACCTGTCATGGGACTATATGCAAAATGAT 143  
QY 537 ctgggtgagctgctctggtggagcgaagaattccctcctcctagatgccaataatttccattc 596  
|||||  
Db 144 CTGGGTGAGCTGCTGTGGACGCAAGATTTCTCTTGAGAGTGCCTAAATCTTTTCCATTC 203  
QY 597 aacacccagtcgtgtacaaagaactctgtgtgaatlcagaactgcttcaccccttcaac 656  
|||||  
Db 204 AACACCTCAGTCCGTACCAAAAGACCTGTGTGTAAATTCACGAGTACACCTTTTCAAC 263  
QY 657 atcgttaagccaaagccaccctggaatggtctgtgtaagtcctacacaccagctgaagtt 716  
|||||  
Db 264 ATCGCTAAGCCAAAGGCCACCCTGGATGGGTGCTGTAGTCTTACCATTCACAGGCTGAGGTT 323  
QY 717 tatgatacagtgatcatlaactaaagaacatggtctcccatcctgtcagttctcagtct 776  
|||||  
Db 324 TATGATAAGTGTGATTAACATTAAGAACATGGCTTCCATCTGTCACTGCTTCATGCT 383  
QY 777 gttgtgtatccactctggaagactctcgaaggagctgaatatgatagcacagaccgtcaa 836  
|||||  
Db 384 GTTGCTATCTCTACTGGAAGAGCTTCTGAGGAGACTAATATGATGTCAGACCAAGTCMA 443  
QY 837 agggagaagaatgataagttcctccctggtggaagcacaatacatatgtctgcgaagtc 896  
|||||  
Db 444 AGGGAGAAAGATGATGATAAGTCTTCCCTGGTGAAGCACCATCATATGCTGCGACAGTGC 503  
QY 897 cttaagaagaatgltccaaatggcctcgcacacactggtccttactcctatctctctc 956  
|||||  
Db 504 CTGAAGAGATGGTCCAAATGGCTGTGACCCACTGTGCTTACTCTCATATCTTCTTCT 563  
QY 957 catgtgacctggttaaaagactgaattcagggcctcatttgagccctcagtaagta 1016  
|||||  
Db 564 CATGTGACCTGATGATAAGACTTGAATTTCAGGCTCATTTGGAGGCCCTACTAGTATGAGA 623  
QY 1017 gaaaggagtcctggccaaagaaagacacagacctgtgcaaatltaactacttltgtc 1076  
|||||  
Db 624 GAAGGAGTCTGGCCAAAGAAAGACACAGACCTTGACCAAAATTTATATCTTTTGGCT 683  
QY 1077 gtatttgataaaggaaagtgtgcactcagaagaacaaagaactccttgatgcaggaatg 1136  
|||||  
Db 684 GTATTGTGATGAAGGAAAGTTGGCACTTCAGAAACAAAGAACTCTTGATGCAAGATAGG 743  
QY 1137 gatgtcactcgtctcgggctggccctaaatgcacacagtcgaatgtltagtaaacagg 1196  
|||||  
Db 744 GATGCTGATCTGCTCGGCTGGCTGCTTAATAATGCACACATGATGTTATGTAACAGG 803  
QY 1197 tctctgcgaagtcgtatgtgatgacacaggaatlcagtcatactgtgcactgtgatatg 1256  
|||||  
Db 804 TCTCTGCAGGCTCGATTGTGATGCAAGAAATCACTATATGCGCATGTGATGGAATG 863  
QY 1257 ggcacacactcctgaagtgcactcaatatctcctcgaaggttcaacacatttctgtgaggaac 1316  
|||||  
Db 864 GGCACCACTCTCTAAGTGCACTCAATATTCCTGAAGGTCAACATTTCTTGTGAGGAAC 923  
QY 1317 catcgcgaagcgtccctcgtgaatctgcacaaactcctcctactcgtgcacaaactctg 1376  
|||||  
Db 924 CATCGCAGGCGTCTTGGAATCTCGGCCAATAATCTTCTTACTGTCATAACACTCTTG 983  
QY 1377 atggaacttggaagttctactgatttgcataatctctccacaaatgataatgagcatg 1436  
|||||  
Db 984 ATGACCTTGGAAGTTCATCTGTTTTCATATCTCTTCCACCAACATGATGAGCATG 1043  
QY 1437 gaagcttatgtcaagaatgagcagctgtccagagaagcccaactaagaatgaagaataat 1496  
|||||  
Db 1044 GAAGCTTATGTCAAGATGACAGCTGTCCAGAGAACCCCACTACCAATGAATAAATAT 1103  
QY 1497 gaagaagcgaagactatgatatgtactcattcattcgaatgatatgtgtcaggttt 1556



|||||  
Db 1104 GAGAAACCGGAGACTATGATGATGATCTACTGATCTGAAATGAGTGTGCTCAGGTTT 1163  
OY 1557 gatgatgaacatctccctcccttataccaattcgctcaatgtgccaagaatcctaa 1616  
Db 1164 GATGATGACAACTCTCTCTTTATCCAAATGCTCTCAGTGGCCAGAGATCCTTAA 1223  
OY 1617 acttgggtacatacatctgccttgaagaaggactggactacgtcccttagtccctc 1676  
Db 1224 ACTTGGGTACATTCATTTGCTGTGAAGAGAGAGACTGGACTATGTCCTTACTGCTC 1283  
OY 1677 gccccgcatgcaagaagttaaaaaatcaatattgaaacaatggccctcagcggaattgt 1736  
Db 1284 GCCCCGATGACAGAGATTAAAGTCAATATTGAAACAAGGCCCTCAGCGGATTGGT 1343  
OY 1737 aggaagtacaaaagctccgaatttaaggcatcacacagatgaaccttaagaactgttaa 1796  
Db 1344 AGGAAGTACAAAAAGTCCGATTTATGGCATACAGATGAAACCTTTAAGACTGCTGA 1403  
OY 1797 gctatcagcatgaatcaggaactcttggaccttacttacttctgggaagtctggacaca 1856  
Db 1404 GCTATTCAGCATGATGATGAGATCTTGGACCTTACTTTATGGGAGTTGGAGACACA 1463  
OY 1857 ctgttgattataattgaatccaagcaagcagacacataaactaccctcaggaatc 1916  
Db 1464 CTGTTGATTATTTAAGAAATCAAGCAAGACCATATTAACATCTACCCTCAGCGAATC 1523  
OY 1917 actgatatcgctcttctgtatccaaggatataccaaggctgtaaaactttgaagt 1976  
Db 1524 ACTGATCTCCCTCTTGTATTCAGAGATTACCAAAAGGTGTAACATTGTGAAGAT 1583  
OY 1977 ttccaactctccaggaagaatactcaataataatgacagctgactgtaagaagtgg 2036  
Db 1584 TTTCCAATTCGCCAGGAGAAATATTCAATATAATGACAGTACGTGTAAGATGGG 1643  
OY 2037 ccaactaatcagaatcctcgtgctgcctgaaccgctactactctagtttcgttaaatgag 2096  
Db 1644 CCAACTAAATCAGATCCTCGGTGCTGACCCGCTAATCTAGTTGCTTAATGTGAG 1703  
OY 2097 agagatctagctcaggaactatctggccctcctcctcactcctacaagaactcttagat 2156  
Db 1704 AGGATCTAGCTTAGAGACTATTTGGCCCTCTCTCCATCTGCTACAAAGAAATCTAGAT 1763  
OY 2157 caaagagaaacagataatgctcagaagaaggaaatgcatcctgttctcgtattgat 2216  
Db 1764 CAAAGAGAAACAGATATGTCAAGACAAAGAGAAATGTCAATCTGTCTGATTTGAT 1823  
OY 2217 gaagaacgaagctgtgtacctacagagaataataaagctttctcccaatccagctgga 2276  
Db 1824 GAGAAACGGAAGCTGTACTCTACAGAGAAATACAAACGCTTCTCCCAATCCAGCTGGA 1883  
OY 2277 gtcagacttgagatcagagattccaagcctccaacatcagcagacatcaatgtcat 2336  
Db 1884 GTGCAGCTTGAGATCCAGAGATTCCAAGCCTCCAACATCATGACAGATCATATGGCAT 1943  
OY 2337 gttcttgatagtttgcagctgtcagttgtttgcatgaggtgacactgtgtacattcta 2396  
Db 1944 GTTTTGTATGTTGTGAGTTGTCAAGTTGTTTGTGATGAGTGGCATCTGATCTGTA 2003  
OY 2397 agcatggagcagagactgtaccttctctgtctctctctgtgataacttcaaacac 2456  
Db 2004 AGATTGTGACACAGACTGACTTCTTCTCTCTCTCTCTCTGTGATATACCTTCAAAAC 2063  
OY 2457 aaaaatgtctatgagacacactcacccattcccatctcagaagaaactgtctctagt 2516  
Db 2064 AAAATGTCATGAAAGACACACTACCTATTTCCATTTCTAGGAGAAATGTCTTCTAG 2123  
OY 2517 tcatgtgaaaaacccaaggttctatggaatctgggtgtgccaacactcagactctcgagaca 2576  
Db 2124 TCGATGAAAAACCCAGGCTATGATTTCTGGGGTCCCAACTCAGACTTTTCGGAACGA 2183  
OY 2577 ggcattgcgccttactgaaagttctagttgtgacaaagaactgtgtatattacagag 2636  
|||||

---

Db 2184 GGCATAGCCGCCCTTACTGAAGGTTTCTAGTTGTGCAAGAAACACTGATATTATACAG 2243  
OY 2637 gacagttatgaagatatttcagcatctgtcgtgagtaaaaaaatgtccattgaacaca 2696  
Db 2244 GACAGTTATGAAGATATTTCAGCATACTTGTGTGATTAATAAATGCCATTGAACCAAGA 2303  
OY 2697 agcttccccaagtattcaagaacaccttagcactagggcaaaagaatttaatgcccacca 2756  
Db 2304 AGCTTCTCCGAATTCAGAGACACCTAGCCTAGGCGCAAAAGCAATTTAATGCCACCCCA 2363  
OY 2757 ccaagtcttgaagcgcatacgaaggaataaactcgttactactctcttgatcagtaagag 2816  
Db 2364 CCAGTCTTGAAACGCCATCAACGGGAAATTAACCTGTAATCTTCACTCTTCAGTCAATCAAG 2423  
OY 2817 gaattgacatgatgatcacatcatcagttgaatgaagaagaagatttgaacttat 2876  
Db 2424 GAAATTGACTATGATGATACCATATCATTTGAATGAAGAAGAAATTTGACATTTAT 2483  
OY 2877 gatgagatgaaaaatcagaagccccgcagcttcaaaaagaacaagacactattat 2936  
Db 2484 GATGAGATGAAAAATCAGAGCCCCCGCAGCTTTCAAAAAGAAACAGACACTATTTAT 2543  
OY 2937 gctgcagtgagagagctctgggattatggaatgagtgctcccaatglttctaagaac 2996  
Db 2544 GCTGCAGTGGAGAGCTCTGGATTATGGATGATGATGATCTCCCATGTTCTAAGAAC 2603  
OY 2997 agggctcagagtgagcagtgctccctcagttcaagaagaatgttttccaagaaattcagat 3056  
Db 2604 AGGGCTCAGAGTGGCAGTGTCTCCCTCAGTTCAAGAAAGTGTGTTTCCAGGAATTTACTGAT 2663  
OY 3057 ggtccttactcagcccttataccgttggagaactaaatgacaatttggagctctggg 3116  
Db 2664 GGCCTCTTACTCAGCCCTTATACCGTGGAGAACTAATGAAACTTTGGAGCTCTGGGG 2723  
OY 3117 ccataataagaagcagaagttggaagataataatcatgtgtaacttccagaatcagcctct 3176  
Db 2724 CCATATATTAAGAGCAAGAGTTGAAGATATATCATATGTAATCTTCAGAAATCAGGCTCT 2783  
OY 3177 gctccctatcccttacttactcagccttacttctatatgaggaagatcagaagcagaagca 3226  
Db 2784 CGTCCCTATTCCTCTATTTAGCCTTATTTCTTATGAGAAATCAGAGGCAAGGACGA 2843  
OY 3237 gaacctagaaaaaacttgtcaagcctaaatgaacaaacttactcttggaaagtgc 3296  
Db 2844 GAACCTAGAAAAAACTTTGTCAAGCCTATATGAACCAAAACTTCTTTGGAAGGTCAA 2903  
OY 3297 catcatatggcaccacaaagaatgagtttgcagcagaagccttggcttattcttctgat 3356  
Db 2904 CATCATATGCAACCCACTAAAGATGATTTGACTGCAAAAGCTGGCTTATTTCTGAT 2963  
OY 3357 gtcgacctggaaaaaatgctgcaatcagccttgaatggaccttctgtctgcacact 3416  
Db 2964 GTTGACCTTGAAAAAGATGTGCACATCAGAGCCTGATTTGGACCCCTTGCTGTCACACT 3023  
OY 3417 aacacactgaacctgtctcatlgygagacaagtgacaatgacagaattgtctgttcttc 3476  
Db 3024 AACACACTGAACCCGTGCTCATAGGAGACAACTGACAGTACAGGAATTTGCTGTTTTTC 3083  
OY 3477 accatcttltgatgacaaaagctgtgacttcaactgaataatggaagaagaactgaag 3536  
Db 3084 ACCATCTTTGATGAGACCAAAAGCTGTGATCTTCACTGAAATATGGAAGAAACCTGACGG 3143  
OY 3537 gtcctctgcaatatccaatggaagatcccaactttaaaggaatlaatcgctcactgca 3596  
Db 3144 GCTCCCTGCATATTCACATATGGAAGATCCCACTTTTAAAGGAATTTATCGTTCCATGCA 3203  
OY 3597 atcaatgtctacataatggaatcactacactgtgcttagtlaatgctcagatcaaaagat 3656  
Db 3204 ATCAATGTGCTACATATATGATACACTACTGCTTAGTAATGAGCTCAGATCAAAAGATT 3263  
OY 3657 cgatgtatctgtctcagatagtgagcagaatgaaanaactcaatctatctcaatcagtgga 3716  
Db 3264 CGATGTATGTGCTCAGATGGCAGCAATGAATAAACAATCATTTCTATTTCAATTCAGTGA 3323  
|||||



QY 3717 catgttctactgtacgaaaaaagagatataaatgacatgttaccatctatcca 3776  
|||||  
Db 3324 CATGTGTTCACTGTRACGAAAAAAGAGAGATATAAATGGCACTGTAAATCTCTATCCA 3383  
QY 3777 ggtgtttttgagacagtgaaatgttaccatccaaagcttggaattggcgggtggaaatgc 3836  
|||||  
Db 3384 GGTGTTTTGAGACAGTGGAAATGTTACCATCCAAAGCTGGAAATTTGGCGGTGGAAATGC 3443  
QY 3837 ctatattggcagcatctacatgtctgggatgagcacactttctgggtgtacaagcaataag 3896  
|||||  
Db 3444 CTATATTGGCGAGCATCTACATAGCTGGGATGAGCACACTTTTTCGTGTACAGCAATTAAG 3503  
QY 3897 tctgacactccctggggaatgtctctgagacatlaagaatttccagattacaagcttca 3956  
|||||  
Db 3504 TGTCAAGACTCCCTTGGAATGGCTTCTGGACACATTAAGATTTTCAGATTACAGCTTCA 3563  
QY 3957 ggaacaatatgacagatgggccccaaagcttggccagacttcatlccggaatcaatcaat 4016  
|||||  
Db 3564 GGACAATATGAGACAGTGGGCCCCAAAGCTGGCCAGACTTCAATTAATCCGGATCAATCAAT 3623  
QY 4017 gcttgaggacacaaaggagccctttcttgatcaagatgagatcgttggcaaccaatgatt 4076  
|||||  
Db 3624 GCTGGAGACCAAGAGACCCCTTTCTTGGATCAAGGTGATCTGTGGCAACCAATGATT 3683  
QY 4077 attcagggcatcaagaccagaggtgcccgttcagaagatctccagctctacatctcag 4136  
|||||  
Db 3684 ATTACAGGATCAAGACCCAGAGGTGCCCTGACAGAGTTCTCAGAGCTCTACATCTCTAG 3743  
QY 4137 ttatcatcatglatagttctgatatggagaagaagatggcagacttaccggaatccact 4196  
|||||  
Db 3744 TTATATCATCATGATGATGCTTGATGGAGAGAGATGGGAGACTTATFCAGGAATTTCCCT 3803  
QY 4197 ggaaccttaatggtctctcttggcaatgtgattcatcttgggataaacaacataatttt 4236  
|||||  
Db 3804 GGAACCTTAAGGTCTTCTTGGCAATGTGATTCATCTGGGATAAACAACAATATTTT 3863  
QY 4257 aacctccaattatgtctcgatatacgttggcaacccaactcatatagatctcgacg 4316  
|||||  
Db 3864 AACCTCCAAATTAATGCTCGATACATCCGTTTGCAACCACTCATTAATAGATTTGCGAGC 3923  
QY 4317 accttcgcatgagatgtagtggcgttgatttaataatgttcgaacatgccaattggaaatg 4376  
|||||  
Db 3924 ACTCTTGCGATGAGATGATGGCTGTGATTAATAATAGTGCAGATGCCATTTGGGAATG 3983  
QY 4377 gagagtaagcaataataatagatgacagagattactgtcttcaatcttacttaacaataattt 4436  
|||||  
Db 3984 GAGAGTAAGCAATATCAGATGACAGATTAATCTGCTCATCTTACCAATATATGTT 4043  
QY 4437 gcaacctggctccttcaaaagctcgacttcaactccaaggaaggaatgactggagga 4496  
|||||  
Db 4044 GCCACCTGGTCTCCTTTCAAAAGCTCGACTTCACTCCAAAGGAGAGATTAATGCTGGAGA 4103  
QY 4497 cctcaggtgaaataatcaaaagatggctgcaagtggaactccgaagaacatgaaatgc 4556  
|||||  
Db 4104 CCTCAGGATTAATATCCAAAGAGTGGCTGCAAGTGCAGCTTCCAGAAAGCATGAAATGC 4163  
QY 4557 acagagatgaactactcaggggagataaatctctgcttaccagcaatgataatgagagttc 4616  
|||||  
Db 4164 ACAGAGATTAATCTCAGGGGAGTAAATCTGCTTACCAACATGATGTAAGAGATTG 4223  
QY 4617 ctcatctccagcagtaagatgagcatcagtcagtcctcttttccagaaatggcaaaatga 4676  
|||||  
Db 4224 CTCATCTCCAGCACTCAAGATGGCCATCAGTGCATCTCTTTTTCAGAAATGGCAAAATGA 4283  
QY 4677 aaggttttccagggaaatcaagactccttcaacactgttggtaactcttlaagaccacg 4736  
|||||  
Db 4284 AAGGTTTTTCAGGGAATCAAGACTCCTTACACCTGTTGGTAAGTCTCTAGACCACCG 4343  
QY 4737 ttactgactgctactccttgaattcaaccccaagattgggtggacacagattggccctgagg 4796  
|||||  
Db 4344 TTACTGACTCGCTACCTTCGAATTCACCCCCAGAGTGGGATGACACAGATTTGCCCTGAGG 4403

QY 4797 atgaggtctcggtcgagagcaagacacttactgactcagcaggttc 4848  
|||||  
Db 4404 ATGAGGTTCTGGCTCGGAGGACACAGCACTTACTAGAGGTGGCCACTGC 4455

RESULT 8  
US-08-882-083-1  
: Sequence 1, Application US/08882083  
: Patent No. 5869292  
: GENERAL INFORMATION:  
: APPLICANT: VOORBERG, Johannes J.  
: TITLE OF INVENTION: HYBRID PROTEINS WITH MODIFIED ACTIVITY  
: NUMBER OF SEQUENCES: 17  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: Foley & Lardner  
: STREET: 3000 K Street, N.W., Suite 500  
: CITY: Washington  
: STATE: D.C.  
: COUNTRY: USA  
: ZIP: 20007-5109  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: Patentin Release #1.0, Version #1.30  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/882,083  
: FILING DATE:  
: CLASSIFICATION: 514  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: US 08/558,107  
: FILING DATE: 13-NOV-1995  
: ATTORNEY/AGENT INFORMATION:  
: NAME: ISACSON, John P.  
: REGISTRATION NUMBER: 33,715  
: REFERENCE/DOCKET NUMBER: 30472/212  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: (202)672-5300  
: TELEFAX: (202)672-5399  
: TELEX: 904136  
: INFORMATION FOR SEO ID NO: 1:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 5035 base pairs  
: TYPE: nucleic acid  
: STRANDEDNESS: single  
: TOPOLOGY: Linear  
: FEATURE:  
: NAME/KEY: CDS  
: LOCATION: 35..5017  
: US-08-882-083-1

Query Match 47.2%; Score 3748; DB 2; Length 5035;  
Best Local Similarity 87.4%; Pred. No. 0;  
Matches 4374; Conservative 0; Mismatches 60; Indels 570; Gaps 2;

QY 415 ccacatgcaaatatgagcttccacactgctcttctgtgaccttggatctgcttta 474  
|||||  
Db 30 CCACCATGGAATATGAGCTTCACACTGCTCTTCTGTGCTTTGGATTTGCTTTA 89  
QY 475 gtgcacacagaagataactactggtgtgagtggaactgtcatggaacatatagcaaatg 534  
|||||  
Db 90 GTGCCACCAAGAAATACTACTGGGTGAGTGGAACTGTCTATGGACTATATGCAAAATG 149  
QY 535 atctcggtgaagctgctgttggacgcaaatcttccctcctagaagtgcacaaatctttccat 594  
|||||  
Db 150 ATCTCGGTGAGCTCGCTGTGGACGCAAGATTTCTCTAGAGTGCAGAAATCTTTTCAT 209  
QY 595 tcaaacctcagctggtgtacaaaaagactcgtttgtagaattcaaggttcaacttttca 654  
|||||  
Db 210 TCAACACCTCAGTGTGTATCAAAAAAGACTGTGTTGTATGAATTCACGGATCACCTTTTCA 269  
QY 655 acatcgctaaagccaagccacctgtagtgggtctgtctagctctacatccagcgtcag 714



Db 270 ACATGCGCTAAGCAAGGCGACCCGTGATGGGTCTGCTAGTGCCTTACCATCCAGGCTGAGG 329  
QY 715 ttctatgacagtggttcattacactaaagacatggtcccatccctggtcattcctcag 774  
Db 330 TTTATGATACAGTGTCTACCTTACCTTAAAGACATGGCTTCCCTGCTGAGTCTTTCATG 389  
QY 775 ctgtctgtctacccacttgaagagctctgagagagcttgaatatgata tcaagaccagtc 834  
Db 390 CTGTTGGTGTATCTTACTAGTGAAGAGCTTCTGAGGAGCTGAATATGATGATCAGACCAGTC 449  
QY 835 aaagggaaagaaagaaatgaataagctctccctcgtgtggaagccaatacatatgtctggcag 894  
Db 450 AAAAGGAAAGAAAGATATAAAGTCTTCCCTGGTGAAGCCATCATATGCTTGGCAGG 509  
QY 895 tccgaaagagaatgttccaatgtgacctgtgaccacctgtgaccttaccctacatctt 954  
Db 510 TCCGTAAGAGATGCTGCTCAATGCTTGCACCTGACCTGCTTACCTTACCTATATCTTT 569  
QY 955 ctcatgtgacctggttaaaagacttgaacttcaagaccttgaagacctactagta tga 1014  
Db 570 CTGATGTGACCTGTGTAAGAGCTTGAANTCAGGCTCATGTGAGCCCTACTAGTATGTA 629  
QY 1015 gagaagggagctctggccaagaaagaaacagaccttgcacaaattataactcttgg 1074  
Db 630 GAGAGGAGAGTCTGGCCAAAGAAAGACACAGACCTTGCACAAATTTATACATCTTTTG 689  
QY 1075 ctgtatttgaagaaagaaagcttggcaactcagaacaaagacctctgtatgacagata 1134  
Db 690 CTGATTTTGAATGAAGGAAAGTGTGGCACTCAGAAACAAAGATCTTGTATGACAGATA 749  
QY 1135 gggagctgtcactcgtctcgggacctggcctaaatgtcacacagtcacatgttaltgtaaca 1194  
Db 750 GGGATGTGCTGATMTCTCTGGGCTGGCCCTAAATGCAACAGTCAATGCTTATGTAACA 809  
QY 1195 ggtctctgcaggtctgtattgtgatgtcccaagaaatcagtcattgtgagctgtgtatgga 1254  
Db 810 GGTCTCTCCAGGCTGTGATGTGATGTCACAGGAAATCACTGATTTGATGATGGA 869  
QY 1255 tgggacaccaccccttgaagtgcacataatctcgaaggtcacacatcttctgtgagga 1314  
Db 870 TGGGACACACCTCTGAATGACCTCAATATCTCTGAAAGTCAACATTTCTTGTGAGGA 929  
QY 1315 accatcgcagagcgctccttgaatctcgcacaaacttcccttaactgtcacaacactct 1374  
Db 930 ACCATGCGCAGGCGCTCTTGGAAATCTGCGCAATTAATCTTCTACTGCTCAACACACTCT 989  
QY 1375 tgaatgaccttggacagtttctactgcttctgcatatctctcccaacaaactgtatgga 1434  
Db 990 TGAATGACCTTGGACAGTTTCTACTGTTTGTCAATATCTTTCACCAACCAATGATGGA 1049  
QY 1435 tggagactatgtcaaaatagaacagctgtccagagaaaccccaactacgaatgaaata 1494  
Db 1050 TGGAGGCTTATGTCAAAATAGACAGCTGTCCAGAGGAACCCCACTAGCAATGMAAAAAA 1109  
QY 1495 atgaagaagcggaagactatgatgtacttactgtatcttgaatgtgagtggtcaggt 1554  
Db 1110 ATGAGAGAGCGGAAGACATATGATGATCTTACTGATTTGAAATGATGATGTGTCAGGT 1169  
QY 1555 ttgatgtgaactctctcccttaccatccaaatcgtctcagttgtccaaagaagatccca 1614  
Db 1170 TTGATGATGACACTCTCTCTCTTATTCCAAAATTCGCTCACTTCCCAAGAAAGATCCGA 1229  
QY 1615 aaacttggatcatatactgtctgtgaagagagactgtgagactatgtcccttaagtc 1674  
Db 1230 AAACCTTGGGTATTTACTTGTCTGTAAGAGAGAGACTGGGACTATGCTCCCTTATGTC 1289  
QY 1675 tgcgccccgatgacaaagtataaaatgcaatattgaaacaaatggccctcagcgatg 1734  
Db 1290 TCGCCCCGATGACAGAGATTATAAAGTCAATATTTGAACAATGGCCCTCAGCGGATGG 1349  
QY 1735 gttaggaagtacaaaaaagctcagattatgtgcatacacagatgaacactttaaagctcgtg 1794  
|||||

---

Db 1350 GTAGGAAGTACAAAAAGTCCGATTTTANGGATFACACAGATGAAACCTTAAAGCTCGTG 1409  
QY 1795 aagcatcagaatgaatcaaggaatccttggagaccttacttataatgggaagcttgaaga 1854  
Db 1410 AAGCTATTGAGATGATAGAGATCTTGAGACCTTTACTTTATGGGAAGTGGAGACA 1469  
QY 1855 caactgttgaatattaaagaaatcaagcagaacacataacatactaccctcagga 1914  
Db 1470 CACTGTTGATTTATTTAAGAAATCAAGCAAGCAGACCATATATCATCTTACCCTCAGGAA 1529  
QY 1915 tcaactgtctcgtctcttgtatccaaggaatccaagaaggtgtataaacatttgaag 1974  
Db 1530 TCACGTATGTCGTCCTTGTATTCAGGAGATTCACAAAAGCTTAAACATTTGAAGG 1589  
QY 1975 atttccaatctgcgcaagagaaatatacaataatgagcaatgtacttgaaga tgg 2034  
Db 1590 ATTTCGAATTTCTGCCAGAGAAATTTCAATATTAATGACACTGACTGATGAGAAATG 1649  
QY 2035 ggcacaactaaatcagatccctcgtcgtgacccgctatactactagtttcgttaata tgg 2094  
Db 1650 GGCCAACTTAATCAGATCCTCGTGCTGACCCGCTATTACTCTAGTTCCGTTAATATGG 1709  
QY 2095 agagagatctagcttcaagacctcatttggccctctcctatctcgtcacaagaagactgag 2154  
Db 1710 AAGAGATCTACCTTCAGGACTCATTTGGCCCTCTCTCATCTGCTACAAAGAAATCTGTAG 1769  
QY 2155 atcaagaagaaacagaataatgctcagacaagaagaaatgtcactcgttcttctgatttg 2214  
Db 1770 ATCAAAAGAGAAACAGATATGTCAGACAAAGAGAAATGTATCTCTGTTCTGTATTTG 1829  
QY 2215 atggaaccggaagctgtgtactcctacagagaataatacaagcttctcccaatccagctg 2274  
Db 1830 ATGAGAACCGAAGCTGTGTACCTCACAAGAAATATCAACGCTTCTCCCAATCCAGCTG 1889  
QY 2275 gagtgcagcttgaagatccaaggttccaagacctccaataltgacagagactcaatggct 2334  
Db 1890 GAGTGAGCTTGAAGATCCAGATTCAGACTCCCAACATCTATGACAGCATATGAGCT 1949  
QY 2335 atgttcttgaatgttgcagttgtcagttgttgcattgattcagatgagtgacatctc 2394  
Db 1950 ATGTTTATATGTTTGCAGTTGTGCATTTGTTTCATGAGGTGTCATACGTGATATTC 2009  
QY 2395 taagcatggagacacagactgtacttctctgtctctctctctctctctctctctctctct 2454  
Db 2010 TAAGATTGAGACACAGACTGTCTTCTGCTTCTGCTTCTGAGATATCTTCAAAC 2069  
QY 2455 acaaaatgtctatgaagacacactcaaccttaccatcttccaagaaagacttcttca 2514  
Db 2070 ACNAAATGCTTATGAGACACACTCACCTATTCCTCATCTTCAGGAAACATGTCTTCA 2129  
QY 2515 tgcgatggaanaaccagagctcctatgtatcttgggtgtgcacaactcagacttcggaaca 2574  
Db 2130 TGTGATGGAANAACCAAGGTATGATTTCTGGTGGCCACAACTCAACATTTGGAAACA 2189  
QY 2575 gaagcatgacgcttacttgaaggttcttagttgtgtcagaagaacactgtgtattatag 2634  
Db 2190 GAGGATGACCCGCTTACTGGAAGGTTTCTAGTTTATTCAGAGGGGAGGAGGACGACG 2249  
QY 2635 aggaagatgaagaatattca-----gcaacttctgtgaaataaca 2679  
Db 2250 ACTATCTGACCTGGAGAGATATTCAGTGAAGACAGACGACTCATCTCAACATCGTCGACA 2309  
QY 2680 atgcattgaacaaagagctctctccgaatctcaagaacaccttagcactagcaaaagc 2739  
Db 2310 GTCGTATTAACCAAGAAAGCTTCTCCAGAAATTAAGACACCTTAGCACTAGGCAAAAGC 2369  
QY 2740 aalttaatgcaac----- 2752  
Db 2370 AATTTAATGCCACCAATTCAGAAATGACATGAGAGACTGACCTGTTTGCAC 2429  
QY 2753 ----- 2752  
Db 2430 ACAGAAACACTATGCTTAAATATCAAAATATGTCCTCTAGTGAATTTGTTGATGCTTGGC 2489



QY 2753 ----- 2752  
Db 2490 GACGAGTCTACTACACATGAGGCTATCCTTATCTGTCTCCAAAGAACCCAAATATGAGA 2549  
QY 2753 ----- 2752  
Db 2550 CTTTTCTGATGATCATCATCTGGAGCAATAGACAGTAATTAACAGCCTGCTGTAATGA 2609  
QY 2753 ----- 2752  
Db 2610 CACACTTCAGGCCACAGCTCCATTCACAGTGGGACATGGTATTTACCCCTGAGTCAGGCC 2669  
QY 2753 ----- 2752  
Db 2670 TCCATTAAATTAATGAAACTGGGGACAACTGCAGATCCTCTTGCTTGGGATACC 2729  
QY 2753 ----- 2752  
Db 2730 ACTGTGTACTCAGATACCAAAAGAGTGAATCCCAAGAGACTACAGAAAAA 2789  
QY 2753 ----- 2752  
Db 2790 CAGCTTTTAAAGAAAAAGATACATTTTGTCCCTGAACGCTGTGAAGCAATCATGCAA 2849  
QY 2753 ----- 2752  
Db 2850 TAGCAGCAATAATGAGGAGCAAAATAAGCCCAATAGAAATCAGCTGGGCCAAAGCAAG 2909  
QY 2753 ----- 2752  
Db 2910 GTAGGACTGAAAGCTGTGCTCTCAAAACCACAGCTTGAAGCGCATCAACGGGAAA 2969  
QY 2785 taactcgtactactcttcacgtcagatcaagaagaatctgctatgatacatatcag 2844  
Db 2970 TAACCTGCTACTCTCTTCACTGATCAAGAGAAATGACTATGATGATACATATACAG 3029  
QY 2845 ttgaatgaagaagaatttgacattatgatgaga tgaanaatcagagcccccga 2904  
Db 3030 TTGAAATGAAGAAAGAAATTTTGACATTTATGATGAGATGAATAACAGCCCCGCA 3089  
QY 2905 gcttcanaaagaacaacgacactatttattgctgcagtgagaggtcttgagattatg 2964  
Db 3090 GCTTTCAAAAGAAACACGACACTATTTTATGTGTGAGAGAGAGGCTCTGGGATTATG 3149  
QY 2965 ggatgagtagctcccaatgctcctaagaacaaggtctcagagtgtccctcagt 3024  
Db 3150 GGATGATGACTCCCCACATGTTCTAAGAAACAGGGCTCAGAGTGGCAGTCCCTCAGT 3209  
QY 3025 tcaagaagaftgtttccagaagaatttactga tggctccttactcagcccttatacgtg 3084  
Db 3210 TCAAGAAAGTTGTTTCCAGGAATTTACTGATGAGCTCTTACTACAGCCCTTATACGGTG 3269  
QY 3085 gagaactaaatgaacatttgagactcctggggccatatataagaagcagaagt tgaagata 3144  
Db 3270 GAGAACTAAATGAACATTTGGGACTCTCGGGGCATATATAAGGCAAGAGTTGAAGATA 3229  
QY 3145 atatcatgtgtaacttcagaatacagggcctctcgtccctattcccttatacttagectta 3204  
Db 3330 ATATCATGCTAACCTTTCAGAAATCAGGCCCTGCTCCATTTCTTATTTAGCCTTGA 3389  
QY 3205 ttctcttatgagagaatcagaagcagaagcagaacctagaanaaactt tgcataagccta 3264  
Db 3390 TTTCTTATGAGGAAGATCAGAGGCAAGGACAAACCTAGAAAAAATTTGTCAAGCCTTA 3449  
QY 3265 atgaanaaccaacttacttcttgaaaagtgcacatcatatgycacccaactaaagt tgaat 3324  
Db 3450 ATGAATAACCAAACTACTTTTGGAAAGTGCATCATATATGCAACCACTAAAGATGAGT 3509  
QY 3325 ttgactgcaaaagcctgggcttattctcttgatgtgactggaanaaagt tgcactcag 3384  
Db 3510 TTGACTGCAAAAGCCTGGCTTATTTCTGTGATGTGACCTGGAAAAAAGATGTGCTCAG 3569

QY 3385 gctctgatacccccctctgctctgcacacataacacactgaaacctgtcattgagagac 3444  
Db 3570 GCGTATTTGAGACCCCTCTGCTGTGGCACACTAACACACTGAACCCGTGCTATGGAGAC 3629  
QY 3445 aagtgacagtaacaggaatttgctctgcttcttcaacalccttggatagaccaaagctggt 3504  
Db 3630 AAGTGACAGTACAGGAATTTGCTGTTTTCACACTCTTTGATGAGACCAAAAGCTGGT 3689  
QY 3505 acttcaactgaaatatgagaagaagaactgcagggctccctcgaatalatccaatggaagatc 3564  
Db 3690 ACTTCACTGAAATATGGAAGAAAGAACTGCAGGGCTCCCTCAATATTCACATGGAAATC 3749  
QY 3565 ccaactttaaagaagaattatcgcttccatgcaatcaatggtctacataatgatacactac 3624  
Db 3750 CCACCTTTAAGAGAAATTTATCGCTTCATGCAATCAATGGCTACATATATGATACACTAC 3809  
QY 3625 ctggcttagtaatggtcagaagatcaaaagatcgatgta tctgctcagcatgagcaga 3684  
Db 3810 CTGGCTTAGTAAATGGCTCAGAGTCAAGAGATGAGTGGTATGTGCTCAGCATGGCAGCA 3869  
QY 3685 atgaanaacatccatctcattcatcttcaatcttcaatgagatggttccactgtacgaaaaaagag 3744  
Db 3870 ATGAATAACATCCATCTTATTCATTTTCAGTGACATGTGCTCACTGTACGAAAAAAGAG 3929  
QY 3745 agtaaaaatgagcactgtacaatctctatccagtggttcttgagacag tggaaatgttac 3804  
Db 3930 AGTATAAATGGCAGCTGTACAAATCTCATCCAGGTGTTTTGAGACAGTGGAAATGTTAC 3889  
QY 3805 catccaaagcttgaatttggcgggtggaatgacctatttggcagacatcatatgcttggga 3864  
Db 3990 CATCCAAGCTGGAATTTGGCGGGTGAATGCCCTTATTTGGCAGACATCTCATGTCTGGGA 4049  
QY 3865 tggagcaacttcttctggtgtacaagcaataagtgtcagactccctgtgggaatggtcttg 3924  
Db 4050 TGAGCACACTTTTCTGTGTGTACAGCAATTAAGTGTGAGCTCCCTGGGAATGGCTTCTG 4109  
QY 3925 gacacatlaagatatcttcagaattacagctcagaagacata tggagccccaagc 3984  
Db 4110 GACACATTAAGATTTTTCAGATTAACAGCTTCAGGACAAATTTGGACAGTGGGCCCAAGC 4169  
QY 3985 tggcagacttcatlaltccggatccaatcaatgaccttggagcacaagaagaccttctct 4044  
Db 4170 TGCCAGACTTCAATTAATTCGGGATCAATCAATGAGCTGTGAGCACCAAGAGACCCCTTTCTT 4229  
QY 4045 ggatcaagtgatggtctgttgacccaatgattatcaaggtcaatgaagcccaaggtgtccc 4104  
Db 4230 GGATCAAGGTGGATCTGTTGGACCAATGATTTATTCAGGCAATCAAGAACCCAGGTTGCC 4289  
QY 4105 gtcagaagttctcagagcctctacatctcagtttcatcatcatgatagtcctgatatgga 4164  
Db 4290 GTCAAGAGTTCTCCAGCCTCTTACATCTCTCAGTTTATCATCATATGATATGCTTGATGGGA 4349  
QY 4165 agaagtggcagacttalcgaagaattccactggaaccttaatggtctctcttggcaatg 4224  
Db 4350 AGAAGTGGCAGACTTATTCGAGGAATTCACACTGGAACCTTAATGAGTCTTCTTGGCAATG 4409  
QY 4225 tggatcatcttggtataaacaacaataattttaacctccaattatgtctgataatcc 4284  
Db 4410 TGGATTCATCTGGGATTAACACAAATATTTTAAACCTCCAAATTAATTTGGCTGATATCC 4469  
QY 4285 gtttgcacccaactcatatagatcatcgacagcactcttgcagaggttga tgggctgtg 4344  
Db 4470 GTTTGCACCCAACTCATTAATAGCATTCGCGACACTCTTCGATGAGATGATGATGGCTGTG 4329  
QY 4345 atttaaatgltgcaagcatgccaattgggaatggaaglaaagaacatcatcagatgycaga 4404  
Db 4530 ATTTAATAATAGTTCAGACATCCATTTGGGAATGAGAGTAAAGCAATATGATGATCAGACA 4589  
QY 4405 ttactgtctatcactacttcaacaatatgtttgcacactggttctccttcaaaagctcgac 4464  
Db 4590 TTACTGCTTCATCTACTTTTACCAATATGTTTGCACCTGGTCTCCTTCAAAAAGCTCGAC 4649  
QY 4465 ttcaactccaagggaggaatgaccttgagacctcagtgaaataatccaagaagtgtgc 4524



|||||  
Db 4650 TTCACCTCCAGGAGAGATGATGCTTGAGACCTCAGGTGAATATCAAAAGAGTGGC 4709  
QY 4525 tgcgaagtgagctcccaagaacatgaaatcaagagtgactactactaaggagtaaat 4584  
Db 4710 TGCAGTGAAGCTTCAGAGCAATGAAAGTCACAGAGTAATCTAGGAGAGAAAT 4769  
QY 4585 ctctgtactacaagatglatgtgaagaggttcctcatalctcaagcagtlcaaatgtgcac 4644  
Db 4770 CTGTGCTTACACGATGATGTGAAGGAGTTCATCATCTCCAGAGTCAGTAAGTGGCCATC 4829  
QY 4645 agtggagctctcttttccaagaatgcaagtaaaagttttccaagggaatacaagactcct 4704  
Db 4830 AGTGAAGCTCTCTTTTTCAGAAATGCAAAAGTAAAGGTTTTCAGGGAATCAAGACTCCT 4889  
QY 4705 tcaaaccttgatgaactctctagaccacgcgttactgactcgcactccttcaatcaac 4764  
Db 4890 TCACACTGTGTAAGCTCTTACACCCAGCTTACTGCTACTGCTTCACTTGAATTCAC 4949  
QY 4765 cccagaagtgtgacacagaatgcccctgagatgagaggttcgtgctgcagagcacag 4824  
Db 4950 CCCAGATTGGGTGACACAGATTGCCCTGAGAGATGAGAGTTCTGGGCTGCGAGGACACAG 5009  
QY 4825 acctctactgactcgagcgagttc 4848  
Db 5010 ACCTCTACTGAGGTGGCCACTGC 5033

RESULT 9  
US-08-558-107-1  
; Sequence 1, Application US/08558107  
; Patent No. 5910481  
; GENERAL INFORMATION:  
; APPLICANT: VOORBERG, Johannes J.  
; TITLE OR INVENTION: HYBRID PROTEINS WITH MODIFIED ACTIVITY  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/558,107  
; FILING DATE: 13-NOV-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: ISACSON, John P.  
; REGISTRATION NUMBER: 33,715  
; REFERENCE/DOCKET NUMBER: 30472/212  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)672-5300  
; TELEFAX: (202)672-5399  
; TELEX: 904136  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 5035 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 35..5017  
; US-08-558-107-1

Query Match 47.2%; Score 3748; DB 2; Length 5035;

Best Local Similarity 87.4%; Pred. No. 0;  
Matches 4374; Conservative 0; Mismatches 60; Indels 570; Gaps 2;  
QY 415 ccacatgcaaatgagagcttccacccgtctcttctgtgctcttgcgattctgcttca 474  
Db 30 CCACCATGAAATAGAGTCTCCACCTGCTCTTCTTGTGTGCTTTTCGATTCGCTTTA 89  
QY 475 gtgcaccacagaagatctactctgtgtgagtgagactgtcattggagctatatacaaatg 534  
Db 90 GTGCACACCAAGAAATACTACTGGGTGACATGGAATGTGATGAGCATATATGCAAGATG 149  
QY 535 atctcgtgtgagctgctgtgtgagcgaagattcctcctctagatggtgccaaatcttccat 594  
Db 150 ATCTCGGTGAGACTGCTGTGTGACCAAAATTTCCCTCTAGAGTGCACAAATCTTTTCCAT 209  
QY 595 tcaaacctcagctgtgtgacaaaagactcgtttgtgaaatgaaggttcacactttca 654  
Db 210 TCACACCTCAGTGTGTACAAAAGACTGTTTGTGAAATTCAGGATCACCTTTTCA 269  
QY 655 acatgcctaaagccaaagccaccctgtgaltggtctgctlaagtlcctacatccaggtgag 714  
Db 270 ACATGCGTAAGCCAAAGCCACCTGAGATGGSTCTGTAGTGCCTACATCCAGGCTGAGG 329  
QY 715 ttatgatacagtggtlcaacttaacacttaagaacatggtctccatcctgtcaagttcatg 774  
Db 330 TTTATGATACAGTGTGTCTACTTAAGAAACATGCGCTTCCATCTGTGACGCTTTCATG 389  
QY 775 ctgtgtgtgtactcctactggaagactctgagagagctgaaatgatgatcaaacagctc 834  
Db 390 CTGTGGGTATCCCTACTGGAAGACTTCTGAGGAGCTGAATATGATGATGACGACGTC 449  
QY 835 aaaggaagaagaatgatatgaatcttccctgtgtggaagccatataatgtctgcagag 894  
Db 450 AAAGGAGAAAGAGATATATAATGCTTCCCTGCTGGAAGCAATATGCTGGCAGG 509  
QY 895 tctgtgaagaagaatggtlccaaatggtccctcgaaccatgtgcttactactatattc 954  
Db 510 TCTGAAAGAGAAATGCTCAATGCGCTCTGACCCACGCTGTGCTTACTACTATCTTT 569  
QY 955 ctcatgtgagactgtgaaagactgaaatcagaccatctgagagcctcctactgattgta 1014  
Db 570 CTCATGTGACCTGTGTAATAACCTTGAATTCAGGCTCATTTGGAGCCCTACTAGTATGTA 629  
QY 1015 gagaaggaagctgtgccaaagaaagacacagactctgtcacaaattatactactttg 1074  
Db 630 GAGAAAGGAGTCTGGCCAAAGAAAGACACAGACCTTGCAAAATTAFACTACTTTTG 689  
QY 1075 ctgtatttgatgaaggaagaaagttgtgcaactcagaacaagaactccttcatcagata 1134  
Db 690 CTGTATTTGATGAAGGGAAGAAAGTTGGCACTCAGAAACAAACAACTCCTGTATGCGAGATA 749  
QY 1135 gggatgtgcactctgtcgtggcctgtggccttaaatgacacagatcatgttatgtaaa 1194  
Db 750 GGGATGCTGCATCTGCTCGGCTGTGGCTTAATAATGCAACAGTAAAGTTATGTAAACA 809  
QY 1195 ggtctctgccaggtctgattgtagtgcacaaagaaatcagttatctgtgcatgtatgtaa 1254  
Db 810 GGTCTCTGCCAGGTCTGTGATTGGATGCCACAGAAATCAGTCTATTGGCATGTGATTGGA 869  
QY 1255 tgggcacacactcgtgaagtgcactcaatatctcctggaaggtcacacattcctgttgaga 1314  
Db 870 TGGGACACCACTCGAAGTGCACATCAATATTCCTCGAAGGTCACACATTTCTGTGAGGA 929  
QY 1315 accatgcgaagcgctccttgaataatctgcgaataacttctactgctccaaacactct 1374  
Db 930 ACCATGCGCAGGCGCTCTGGAAATCTGCGCAATTAACCTTCTTACGTCTCAAAACACTCT 989  
QY 1375 tgalgaccttggacagtttctactgattgtcatatctctccacacacatgatgca 1434  
Db 990 TGATGAGACTTGTGACAGTTTCTACTGTTTGTCTATATCTCTTCCACCAACATGATGGA 1049  
QY 1435 tggaaagcttatgtcaaaagtagacagctgtccagaggaaccccaactacgaatgaaata 1494  
|||||



Db	1050	TGGAAGCTTATGTTAAAGTGACAGCTGTGCCAGAGAACCCCACTACGATGAAAATA	1109
Qy	1495	atgaagaagcggaaagactatgatgatccttcaacttcaatctbaaitygctcagct	1554
Db	1110	ATGAAGAAGCGGAAGACTATGATGATGATGATCTTCTGATCTCAATFGATGTGTCAGCT	1169
Qy	1555	ttgatgatgacaactctccttcccttcaatctgcctcagctgtgccaaagaactccta	1614
Db	1170	TTGATGATGACACACTCTCTCTCTTATCCAAATTCGCTCACTGTCCAAGAAGCATCTCTA	1229
Qy	1615	aaactgtgtacatctacaactgtcctcgtgaagaaggagacgtggagcatgtctcccttagcc	1674
Db	1230	AAACTTGAGTCACTTACATTTGCTGCTGAAAGGAGGAGCTGGGACATATGCTCTCCCTTAGTCC	1289
Qy	1675	tcgcccccgatgacaagaagttataaagaatcaatttgacacatgtgcctcagcggatgtg	1734
Db	1290	TGCGCCCCGATGACAGAAAGTTATTAAGAATCAATTTTGAACAATGGCCCTCAGCGGATG	1349
Qy	1735	gtaggaagtlacaaaaaagtlccgattatgtgcatacagatgaacaccttaagactcgtg	1794
Db	1350	GTAGGAGGTACAAAAAAGTCCCGATTATGCGATACACAGATGAACCTTTAAAGACTCGTG	1409
Qy	1795	aagctatcgcagcatgaatcsgaatcttggagccttacttatabgggaaagtltgagaca	1854
Db	1410	AAGCTATTTCAGCATGAATTCAGGAATTTGGAGACTTTACTTTATGGAAGTGTGAGCA	1469
Qy	1855	caacttgattatataatttaagaatccaagaagcagaccataacaatcacctcaagaa	1914
Db	1470	CACGTGTGATTATATTTTAAATATCAAGCAAGCAAGCAATATATACCTTACCTTCAGGAA	1529
Qy	1915	tcaactgatgtccgtcctcttgattatccaaggagattacaaaagtgltaaacatltgaaag	1974
Db	1530	TCACGATGTGCCGCTCTTTGTATTTCAAAGGAGATTACAAAAGGTGTAAACATTTGAAAG	1589
Qy	1975	atttccaattctgcgcaggagaaatatctcaatatataatgycagctgtacgttagaagatg	2034
Db	1590	ATTTTCCAAATTTCTCCACGGAGAAATATTCAAAATATAAGCAAGTACTGTAGAAAGTG	1649
Qy	2035	ggccaactaaatcagatccctcgtgtgcctcagccgcatctaacctagttctgttaataag	2094
Db	1650	GGCCACTAATATCAGATCCCTGGGTGCTGACCCGCGATTACTGTAGTTTCGTTAATATGG	1709
Qy	2095	agagagatctagcttcaagacatcatgtgcccctctccatactcgtacaaaagaatctgag	2154
Db	1710	AGAGAGATCTAGCTTCAGGACTCATTTGGCCCTCTCCTCATCTGCTACAAAGATCTGAG	1769
Qy	2155	atcaagaaggaaaccagataatgttccaagaagaagatgtcatcctgtttctgtatltg	2214
Db	1770	ATCAAAAGAGAAACCAATATATGTCTGAGAACAGGAGATGTCAATCTGTTTCTGTATATGG	1829
Qy	2215	atggaagacgaagctgtgtacatccagagaatatacaagcttcttcccacatccagtg	2274
Db	1830	ATGGAAGCCGAAAGCTGGTACTCTACAGAGAAATATACAGCGCTTTTCTCCCAATCCAGCTG	1889
Qy	2275	gaatgtcagctttsagatccaagagtlccaagcctccaacatcatgtcacagcatcaatgtgt	2334
Db	1890	GAGTGCAGCTTGAGAGATCCAGAGTTCCAGACCTCCCAACATCATGTCAGACGATCAATGGCT	1949
Qy	2335	atgtcttttgtagtttgcgagtgctgtcagttgttgttcagatgaggtgtgcatactcgtfacatct	2394
Db	1950	ATGTTTTTGTATGATTTTGCAATTTGTCAGTTTGTGTTTCATCAGAGTGGCAATACGTGTACATTC	2009
Qy	2395	taagcatcttgagacaagacagacttccctcctcgtctcctctcctcgtatataccttcaac	2454
Db	2010	TAAACATTGAGACACAGACTGACTTCTCTTCTGTCTTCTTCTCTCTGATATACCTTCAAAC	2069
Qy	2455	acaaatgtgtctatlgaaacacacatccactatccocatcttcagagaaactgtcttca	2514
Db	2070	ACAAAAATGTCATATGAAGACACATCAACCTATPCCCAATTCACAGAAACGTCTCTTCA	2129
Qy	2515	ttctcgtatggaaaaaccagtgctatgtatcttcygggtgtccacaactcagacttctcgaca	2574
Db	2130	TGTGTGATGGAAAAACCAAGGCTCATATGATATCTGGGGTGGCCACAACTCAACATCTTGGGAA	2189

QY	2575	gaagcatabacccgcttacttgaagglttctagtgtgcaagaacactgltatcaag	2633
Db	2190	GAGGATCATCCGCCCTTACTGTGAAGGTTCTGTGTGATTTCCAGGGGGGAGGAGCGACG	2249
QY	2635	aggacagttatgaagatatltcca-----gcatctgcgcgagtaaaaca	2679
Db	2250	ACTATCTGGACCTGGAGAGATATTCTCAGTGAAGACGACACTTCATTCGACATCTGCGACA	2309
QY	2680	atgcacatgaaaccaagaagcttctccagaatccaagaacaccctagcacatgagcaaaagc	2735
Db	2310	GTCGTATTGAAACCAAGAAAGCTTCTCCAGAAATTCAGACACCCCTAGCATTAGCAAAAAGC	2369
QY	2740	aatttaagccac-----	2753
Db	2370	AATTTAATGCCACCAACAATTCGAGAAAATGACATAGAGAAGACTGACCCCTGGTTTGAC	2429
QY	2753	-----	2752
Db	2430	ACAGAACACCTATGCGTAAATAATACAAAATGTCTCCTCTAGTGATTTGTTGATGCTGTGC	2489
QY	2753	-----	2752
Db	2490	GACAGAGTCTACTCCACATGGGCTATCTTATCTGATCTCCAGAAAGCAAAATATGAGA	2549
QY	2753	-----	2752
Db	2550	CTTTTTCGTGATGATCCATCACCTGGAGCAATAGACAGTAATAAACAGCCGTCTGAAATGA	2609
QY	2753	-----	2752
Db	2610	CACACTTCAGGCCACACCTCCATCCACAGTGGGACATGGATTTACCCCTGAGTCAGGC	2669
QY	2753	-----	2752
Db	2670	TCCAAATTAGATTAAATGAGAACTGGGGCAACTGCGAGATCCTCTTGCTGGGATACC	2729
QY	2753	-----	2752
Db	2730	ACTATGTGACTCAGATATCCAAAAGAGAGTGGAAATCCCAAGAAAGTACCAAGAAAAA	2789
QY	2753	-----	2752
Db	2790	CAGCTTTTAAAGAAAAAGATACATTTTGTCCCTGAGACGCTTGTAAGCAATCATGCA	2849
QY	2753	-----	2752
Db	2850	TAGCAGCAATTAATGAGGGACAAATAAGCCCAATTAGAAGTCACCTGGGCAAGACAG	2909
QY	2753	-----ccacacagltctgaaacgcaatcaacgggaaa	2784
Db	2910	GTAGAGACTGAAGGCTGTGCTCTCAAAAACCCACCACTCTTGAACGCCATCAACGGGAAA	2969
QY	2785	taactcgtactactcttcagatcaagaagggaaaattactatgatatataaccatacag	2844
Db	2970	TAACTCGTACTACTCTTCAGTCAGATCAGAGAAATGACTATGATGTATACCATATATAG	3029
QY	2845	ctgaatgaaagaagaagattctgacattatgatgataagatgaaatcagagcccccga	2904
Db	3030	TTGAAATGAGAAAGAGAGATTTTGACATTTTATGATGAGGATGAAATAATCAGACCCGCCA	3089
QY	2905	gcttctcaaaagaacaacagcaactatttatatgctgagtggaagaggtcttggagatatg	2964
Db	3090	GCTTTTCAAAAGAAAACGACGACACTATTTTAAATTCGTCGACGTGAGAGGCTCTGGGATTTATG	3149
QY	2965	ggaatgagtagctcccccacatgctctcaagaacaaggtctcagagtggagagtgtccctcagt	3024
Db	3150	GGATGAGTAGCTCCCCCACTGTTCTTAAGAAACAGGGCTTAGAGTGGCACTGTCCCTCAGT	3209
QY	3025	tcaagaagtglttctccaggaattactatgatgagtccttactcagccctataccgtg	3084
Db	3210	TCAAGAAAGTGTGTTTCCAGGAATTTACTGATGGCTCTTACTCAGCCCTTATATCCGTG	3269











QY 2215 atgagaacgaagctggtacctacacagagaataatacaagcttctcccaatcaagctg 2274  
 |||||||  
 Db 1830 ATGGAACCGAAGCTGGTACCTCACAGAGAAATATCAAGCTTTCTCCCAATCCAGCTG 1889  
 QY 2275 gagcgagctgagagatcacagagttccaagcctccaacatcatgacagatcaatggtc 2334  
 |||||||  
 Db 1890 GAGTGCACCTTGAGGATCCAGAGTTCCAAAGCTCCACATCTCATGCAAGCATCAATGGCT 1949  
 QY 2335 atgtttttagatgttgcagttgtcagttgttgcagttgttgcagttgttgcagttgttgc 2394  
 |||||||  
 Db 1950 ATGTTTGTAGTATGTTGAGTTGAGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2009  
 QY 2395 taagcaattggagacagacagctgaactccttctgtcttcttcttcttcttcttcttct 2454  
 |||||||  
 Db 2010 TAAACATTTGGAGCAGACAGCTACTCTTCTGCTCTTCTGCTCTGATATATACCTTCAAC 2069  
 QY 2455 acaaaatggtatagaagaacacactcaccctatcccatcttccaagaagaactgtcttca 2514  
 |||||||  
 Db 2070 ACNAATGTTGTTATGTAAGACACACTACCTATTCCTCATCTCAGAGAACTGTCTTCA 2129  
 QY 2515 tgcgatggaaaacccaggtctatagatcttgggggtgcacacacacagacttcggaaca 2574  
 |||||||  
 Db 2130 TGTGATGGAAACCCAGAGTGTATGATTTCTGGGGTGCACAACTCAGACTTTGGGAACA 2189  
 QY 2575 gaggcataagccgcttactgaagttcttgaattgttgcagaagaacatggtatctacg 2634  
 |||||||  
 Db 2190 GAGGCAATGACCGCTTACTGTAAGGTTTCTAGTTTCTTCCAGAGGGGGAGGAGCAGACG 2249  
 QY 2635 aggcagttatagaagatcttca-----gcatcttgcctgagtaaaaca 2679  
 |||||||  
 Db 2250 ACTATCTGGACCTGGAGAGATATTCAGTGAAGACGACGACTACATCCAGATCGTCGACA 2309  
 QY 2680 atgcatttgaacccaagaagcttctccagaattccaagaacacactagacatagcaaaaag 2739  
 |||||||  
 Db 2310 GTCTGATTGAACCAAGAGCTTCTCCAGAAATTCAGACACCTCTAGCAAAAGC 2369  
 QY 2740 aatttaatgcaac----- 2752  
 |||||||  
 Db 2370 AATTTAATGCCACACAAATTCAGAAATGACATAGAGAAGCTGACCTGGTTTGCAC 2429  
 QY 2753 ----- 2752  
 Db 2430 ACAGAACACCTATGCTTAAATACAAATGTCTCTAGTATTTGTTGATGCTCTTGC 2489  
 QY 2753 ----- 2752  
 Db 2490 GACAGAGTCTACTTCACATGGGCTATCTTATCTGATCTCCAGAGCAAAATATGAGA 2549  
 QY 2753 ----- 2752  
 Db 2550 CTTTCTTGATGATCATCACCTGGAGCAATAGACAGTAAATACAGCCTGTCTGAATGA 2609  
 QY 2753 ----- 2752  
 Db 2610 CACACTTCAGGCCACAGCTCCATCACAGTGGGAGCATGTATTATCCCTGAGTAGGCC 2669  
 QY 2753 ----- 2752  
 Db 2670 TCCAAATTAAGATTAAATGAGAACTGGGGACAACAGATGCTTGTGCTGGATTAAC 2729  
 QY 2753 ----- 2752  
 Db 2730 ACTATGTACTCAGATACCAAAAGAGAGTAAATCCCAAGAGAGTCCACGAGAAAAA 2789  
 QY 2753 ----- 2752  
 Db 2790 CAGCTTTAAGAAAAAGATACCATTTTGTCCCTGAGCGTTTGTGAAGCAATCATGCAA 2849  
 QY 2753 ----- 2752  
 Db 2850 TAGCAGCAATTAATGAGGAGCAAAATAAGCCGAATAAGAGTACCTGGGCAAGCAAG 2909  
 QY 2753 -----cccaacagctcttgaacgccaatcaacgggaaa 2784

Db 2910 GTAGACTGAAAGGCTGTGCTCTCAAAACCCACCAAGTCTTGAAGCCCATCAAGGGAAA 2969  
 QY 2785 taactgtaactactctcaagtcagatcaagaaggaattgacatgatatacatcaag 2844  
 |||||||  
 Db 2970 TAACTCGTACTCTCTTCAAGTCAAGAGGAATTTGATATGATATGATATCAATATCAG 3029  
 QY 2845 ttgaatgaagaagaagatttgacattatgatgatgatgatgatgatgatgatgatgatgat 2904  
 |||||||  
 Db 3030 TTGAATGAAAGAGAGAGATTTTGACATTTATGATGAGAGATGAATAATCAGAGCCCGCA 3089  
 QY 2905 gcttccaaagaagaacagacacttattatgtcgcagtgagaggtctcggattatg 2964  
 |||||||  
 Db 3090 GCTTTCAAAAGAAACAGACACTATTTTATGCTGAGTGAAGGCTCTGGGATTAAG 3149  
 QY 2965 ggaatgagtcctccacagtgcttcaagaagaagggcgaagtggcagtgctcccaag 3024  
 |||||||  
 Db 3150 GGATAGTATGCTCCCAATGTTCTTAAGAAACAGGGCTCAAGATGCGATCCCTCAGT 3209  
 QY 3025 tcaagaagttgtttccaagaattactgaatggctccttactcaagccctataccgtg 3084  
 |||||||  
 Db 3210 TCAAGAAAGTTGTTTCCAGGAATTTACTGATGGCTCTTACTCAGAGCTTATACCGCTG 3269  
 QY 3085 gagaactaaatgaacatttgggactcctggggccatatataagaagaagttgaagata 3144  
 |||||||  
 Db 3270 GAGAACTAAATGAACTTTGGAGCTCTGGGGCATATATTAAGCAGAGAACTTGAAGATA 3329  
 QY 3145 atatcatgtiaacttcaagaatcaagcctctgcctcctatcttcttcttcttcttcttct 3204  
 |||||||  
 Db 3330 ATATATGTGTAATCTTCAAAATCAAGGCTCTGCTGCTTCTTCTTATTTCTAGCTTGA 3389  
 QY 3205 ttcttataagaagaatcaagaagcagaagaacttagaaaaacttcttcaagctta 3264  
 |||||||  
 Db 3390 TTTCTTATGAGGAAGATCAGAGGCAAGGAGCAAGCACTTGAAGAACTTGTCAAGCTTA 3449  
 QY 3265 atgaacccaacttacttlttggaaagtgcacatcatatgacccactaaagaatgagt 3324  
 |||||||  
 Db 3450 ATGAACCAAAACCTTATGTTGGAAAGTGCAACATCATATGACCCCACTAAATATAGT 3509  
 QY 3325 ttgactgcaaaagcctgggcttattctctgtatgttgcacttgaagaagaatgtgcacag 3384  
 |||||||  
 Db 3510 TTGACTGCAAAAGCCTGGGCTTATTTCTGATGTTGACCTGGAAGAAAGATGTGACACTCAG 3569  
 QY 3385 gccgtatggagccctctgtctgcacactaaacacatgaacccgtgctcatagagagac 3444  
 |||||||  
 Db 3570 GCCTGATTGGACCCCTCTGCTGCTCCACACTAACACACTGAACCCCTGCTATGGAGAC 3629  
 QY 3445 aagtgacagtcagaatltgtctgttcttccacatcttggatgaagacccaagctggtc 3504  
 |||||||  
 Db 3630 AAGTGACAGTACAGGAATTTGCTGTTTTCACCATCTTGTGATGAGAGCCAAAGCTGTGT 3689  
 QY 3505 acttcaactgaataatgtgaagaagaactgcagggctccctgcaatatcagaatgaagatc 3564  
 |||||||  
 Db 3690 ACTTACTGAAATATGTGAAGAAAGAACTCAGGGCTCCCTGCAATATTCAGATGGAATATC 3749  
 QY 3565 ccaactttaagaagaatatacgtctcatalgcaatcaatgylacataatgataactac 3624  
 |||||||  
 Db 3750 CCACCTTTAAAGAGAAATATGCTTCCATGCAATCAATGAGTATGATGATGATGATGATGAT 3809  
 QY 3625 ctggcttagtaatggtcagaatgaagaatgtgaatgtatctgtcgaagatggcgagca 3664  
 |||||||  
 Db 3810 CTGGCTTAGTATGCTCAGATCAAAAGATTGATGATGATGATGATGATGATGATGATGATG 3869  
 QY 3685 atgaacaatcatctatctatcttcaagtgagatgttcaactgtatcgaagaagaagaag 3744  
 |||||||  
 Db 3870 ATGAATAATCATCTATTTCTATTTCACTGAGCACTGTGCTGATGCAAGAAAAAGAGG 3929  
 QY 3745 agtatataatggcactgtacaaatctatccaaagtgctttagacagtggaatgttact 3804  
 |||||||  
 Db 3930 AGTATTAATATGCACTGTACATCTCTATCCAGGTGTTTGAAGACAGTGAATGTATAC 3989  
 QY 3805 catccaagctggaaattggcggttggaatgcttatttggcgagcatctatcatgcttggga 3864  
 |||||||



Db 3990 CATCAAGCTGGAAATTTGGCGGGTGGAAATGCTTATTTGGCGAGCATCTCATGTGGGA 4049  
QY 3865 tggagcaactttcttgggtgtacagaataaagtgtaagctccccgggaatggtcttg 3924  
Db 4050 TGAAGCACTTTTCTGGTGTACAGCAATAGTGTAGACTCCCTGGGAATGGCTTCTG 4109  
QY 3925 gacacattagagatttcagattacagctcgaagacaataatggaagctgggccccaaagc 3984  
Db 4110 GAACATTAAGATTTTTCAGATTACAGCTTCAGAGACAATATGACAGTGGGCCCAAGC 4169  
QY 3985 tggcgaacttcattatccgatacaataaagtcctgagagaccagaagcccttctt 4044  
Db 4170 TGGCCAGACTTCTATTTCCGGATTCATCAATGCTGAGACCAAGAGACCCTTTTCTT 4229  
QY 4045 ggaataagttgagctgttggcccaatgatatattacaggaatcaagaccaggggtccc 4104  
Db 4230 GGATCAAGGTGGATGTTGGTGGACCAATATTTTACGGCATCAAGACCAAGGGTGGCC 4289  
QY 4105 gtcaagaattctccagactacatctcagtttatacatcagtaagttcttgatggga 4164  
Db 4290 GTCAAGAGTTCTCCAGCTCTACATCTCTCAGTTTATCATCATGTATAGCTTGATGGGA 4349  
QY 4165 agaagtggaagactatcgaagaaattccactggaaccttaatggtctctcttggaatg 4224  
Db 4350 AGAAGTGGCAGACTTATTCAGAGAAATTCACCTGGAACCTTAATGCTTCTTGGCAATG 4409  
QY 4225 tggattcatctgggataaaacaataattttaaaccctccaattatgtctgaatcc 4284  
Db 4410 TGGATTCATCTGGATTAACACAAATATTTTAAACCTCCAAATATTTGCTGATACATCC 4469  
QY 4285 gtttgaccacactatlaagatcgcagacactcttcgcatggagtggatggctgtg 4344  
Db 4470 GTTTCACCACTCTTATATAGCATTCGACGACTCTTGCAATGGAGTTGATGGCTGTG 4529  
QY 4345 atttaaatgtgcaagatgcatctgggaatggagaatgaagaatatacagatgcacaga 4404  
Db 4530 ATTTAAATAGTTTCAGCATATGCCATTTGGGAATGAGAGTAAGCAATATCATGATGCACAGA 4589  
QY 4405 ttactgtctcatctacttaccataatgtttgcaacctggtctccttcaaaagctcgac 4464  
Db 4590 TTAAGCTTTCATCTTACTTATCCAAATATGTTTCCACCTGGTCTCTTCAAAAAGCTCGAC 4649  
QY 4465 ttcaactccaaggaagaaatgaatgctggaagactcaggtgaaataatccaagaatggc 4524  
Db 4650 TTTCACCTCCAGGAGAGATGAATGCTTGAGACTCAGAGATTAATCCAAAAGATGGC 4709  
QY 4525 tgcgaatggaactccagaagacaaatgaatgcaaggaatgaactcagggagataaat 4584  
Db 4710 TGCAGTGGACTTCCAGAGACAAATGAATGACAGAGTAAGTCAAGGAGTAAAT 4769  
QY 4585 ctctgtctccagcatgatatgaaagagttctctatctccagcaatcaagatggccatc 4644  
Db 4770 CTCTGCTTACCAACATGATGTGAAGAGTCTCTCATCTCCACCACTCAAGATGGCCATC 4829  
QY 4645 agtggactctcttttccagaatggaagaatggttttccagggaaatcaagaactcct 4704  
Db 4830 AGTGGACTCTCTTTTTCAGAAATGGAAGTAAGSTTTTTCAGGAAATCAAGACTCTCT 4889  
QY 4705 tcaacactgtgtgaactctctagaccacagcttactgactcgtaactcctgaattcac 4764  
Db 4890 TCACACCTGTGGTGAAGCTCTGTAGACCAACCGTTACTGACTCGCTTCAATTCAGATTCAC 4949  
QY 4765 ccagaagttgggtgacacgaatgctgctgaagatggaagttctggctcgaagacag 4824  
Db 4950 CCCAGAGTTGGGTGACACAGATTTGCCCTGAGAGTGTCTGGCTGCGAGGACAG 5009  
QY 4825 acctactgactcgaagagttc 4848  
Db 5010 ACCTTACTGAGGGTGGCCACTGC 5033

RESULT 11  
US-08-670-707A-38

; Sequence 38, Application US/08670707A  
; Patent No. 5859204  
; GENERAL INFORMATION:  
; APPLICANT: Lollar, John S.  
; TITLE OF INVENTION: Hybrid Human/Animal Factor VIII  
; NUMBER OF SEQUENCES: 40  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.  
; STREET: 5370 Manhattan Circle Suite 201  
; CITY: Boulder  
; STATE: Colorado  
; COUNTRY: USA  
; ZIP: 80303  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/670,707A  
; FILING DATE: 26-JUN-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: WO PCT/US94/13200  
; FILING DATE: 15-NOV-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/212,133  
; FILING DATE: 11-MAR-1994  
; APPLICATION NUMBER: US 07/864,004  
; FILING DATE: 07-APR-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Greenlee, Lorraine L.  
; REGISTRATION NUMBER: 27,894  
; REFERENCE/DOCKET NUMBER: 75-95F  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 303/499-8080  
; TELEFAX: 303/499-8089  
; INFORMATION FOR SEQ ID NO: 38:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4334 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: not relevant  
; MOLECULE TYPE: cDNA to mRNA  
; HYPOTHETICAL: NO  
; ORIGINAL SOURCE:  
; INDIVIDUAL ISOLATE: Factor VIII lacking B domain  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 3..4334  
; US-08-670-707A-38  
  
Query Match 40.2%; Score 3189.8; DB 2; Length 4334;  
Best Local Similarity 83.5%; Pred. No. 0;  
Matches 3692; Conservative 0; Mismatches 637; Indels 90; Gaps 3;  
  
QY 420 atgcaaatagagctccacactgctcttctctgtgctcgttggaattcgttaagtc 479  
Db 3 ATCCACTGAGAGCTCCACACTGTGCTTCTGTGTCTGTGCTGTGCACTCGGCTTATGCGC 62  
QY 480 accagaagataactcctgggtgagtggaactgcatggagactatatgcaaatga---t 536  
Db 63 ATCAGAGATATCACTGCGGCGCAGTGAAGTCTCTGAGACTACCGGCAAGTGAATC 122  
QY 537 ctctggagactgctgtggaacgaagaattccctcctgaagtggaagaaacttttcacatc 596  
Db 123 CTCCTGAGCTGCACGTGAGCACACAGATTTCTGCTTACAGCGCCAGAGACTCTTCCGTTG 182  
QY 597 aacactcagctggtfacaaaagactcgttgtagaattcaagggttaaccctttcaac 656  
Db 183 GGCCTCACTGCTCTGTACAAAAGACTGTGTCTGTAGAGTTACGAGTATCACTTTTCAGC 242



QY	657	atcgcataagccaaagccaccctcggatcgtcgtcagtcctaccacccagctcgaagtc	716
Db	243	gtttgccagggcccgagccaccatggatggctgctcgtggcttccatccacccagggctt	302
QY	717	tatgataacagttgltcaatctaacctaagaatcgtcccatccctcgtcagttcttcagtc	776
Db	303	ttacgacacggctgctgttacccttgaaamaaacatggctttctatcccggttagcttcacgct	362
QY	777	gtctgctgtaaccaccccgaaagctcttcgagggagctgaatacgtatgaacagcagtc	836
Db	363	gtcgggctctcctcttcttggaatcttccgaagggcgcctaatatagatgacacccagccaa	422
QY	837	agggaagaaagaaatgaataagcttcctccctcgtgtgaaagcatatactctcgcagtc	896
Db	423	agggaagaaagaaacacatataagctcttcccggttaaaagccaaacctgcgtgcagagtc	482
QY	897	ctgaaagagaatcgttccaatcgtcctcgcaccacctcgtgccttacctatactcttc	956
Db	483	ctgaaagaaatgctctccacacgctctcgaccacacatgctctcaccttctcttacctgct	542
QY	957	catctgaccttggtaaaagacttgaattcgaagctcctaattgagccctactagatgta	1016
Db	543	cacgtgacctggtgaaagacgtgaatttcgggctctatgtagagccctgctggtttgtaga	602
QY	1017	gaaggaagctcgtgcacagaagaaagacaagaccttcgacaataattactacttltgct	1076
Db	603	gaaggaagctctgacccagaaagaaagaccacacacctgcacgaatttgatcatttttgc	662
QY	1077	gataattgaagaaggaaagattcggcactcgaagaaacaaagacctcttgatcgaagatag	1136
Db	663	gtctttgatgaaaggaaatttgccacttcagcagaaatgactctcttgacacgggcatg	722
QY	1137	gatctgcatactcgtcgggctcgtgcctaaatgcaacagtcacatggtltaatgtaacag	1196
Db	723	gatcccgacacctgcagggccacgacctgacatgcacacagtcattggctatgtaacagg	782
QY	1197	tcctcgtccagttctgatttggaatgccacagaatcagttctatcgtcattgattgtaatg	1256
Db	783	tcctctccacaggtctgattgcgatgcatatgaaamaaacgtctactgacagctgattggatg	842
QY	1257	ggcaccaacctcctaagtgcactaaatctctcgaaggtaacaaactctctgtgaagaaac	1316
Db	843	ggcacacagcccggaagtgcactccatttttcttgaaagccacacgitttctgtagggcac	902
QY	1317	catcgcagggcgtcctctggaatctgcgcaataactctccttactcgtcgaacaactcgtg	1376
Db	903	catccgacagccttcccttgagatctgcgcacattcttccgacctgctcagacacatfctcg	962
QY	1377	atggaccttggagaagattctcactggttltgtcatatctctctcccaacaatgatggcagt	1436
Db	963	atggaccttggccagatttccactggtttgtcatatcttctccacacacatagtgggatg	1022
QY	1437	gaagctatgtcaaatgtagacgtctgcgaaggaaaccccaactagaatgaanaataat	1496
Db	1023	gaagctcactcgtcagatgaaactgcggccgagagagccccaagctcggagaaagctgatt	1082
QY	1497	gaagaagcgaagaactatgatactctactcgtatctgaatgtagtctgtaagttc	1556
Db	1083	gaaga---ggaaatattatgatacaatttggatgcacatgcgacatgagcagtgctgcgctc	1139
QY	1557	gatgatagacaactcctcctcttataccaatctgcctcagttgcgaagaagcatctcaaa	1616
Db	1140	gatgatagacacagctgcttcccttttatcccaatccgctgctgttgccaaagaaacatccaaa	1199
QY	1617	acttgggtatacattacatctgctcgtcgaaggaaggaactcgtccctcttagctc	1676
Db	1200	acctgggtgactactactctctgcagagagagagacttgggactacgccccggcgctccc	1259
QY	1677	ggccccgatgcagaagatttaaaagtcaalatttgacaatcgtccctcagcggatgatt	1736
Db	1260	agccccagtgacagaagaatttataaaagctctgactttgaaacagtggctctcagcgaattggt	1319

QY	1737	aggaagtcacaaaagcgcgattatbgctacacagatgaaacotttgaagctcgtgaa	1790
Db	1320	AGGAATAATCAAAAAAGCTCGATTCTGCTGTACCGGATGTAAACATTAAAGCTGTAA	1379
QY	1797	gctatcagctctgaacgaagaaactctggagaccttactttaggggaagtctggagaca	1856
Db	1380	GCTATTCCGTTGTGANTAGGAATCTGGACCTTTACTTTATGGAAAGTTGGAGACAA	1439
QY	1857	ctgtgatataatlaagaaatcaagcaagaagacaaatalaaactctacccctcaagaa	1916
Db	1440	CTTTTGATTATATTAAAGATAAAGGACGCGACCATATATACATCCATCCATGAAATC	1499
QY	1917	actgatctcgtccctctgttatcaaggagattaccaaagctgtaaaacatttgaagt	1976
Db	1500	ACTGATGTCACCGCTTTTGCAACCCAGGAGACTTCTAAAGTTGGAAACATTTGAAG	1559
QY	1977	ttccaattctgcagaggaataatctcaaatlaaatggacagtgcctgaagaatgg	2036
Db	1560	ATGCCAATTCTGCCAGAGAGACTTTCAATATAATGGACAGTACTGTGGAAATGG	1619
QY	2037	ccaataaatacagatccctcgtgcctgaccccgatctactagtctcgttaatatgg	2096
Db	1620	CCAACCAAGTCCGATCCTCGGTGCTGACCCGCTACTACTCGAGCTCCCTTAATCTA	1679
QY	2097	agagatctagcttaagacatccattggccctctctcatctgctacaagaatctgta	2156
Db	1680	AAAGTCCTGGCTTTGGGACATCATTTGGCCCTCTCATCTGCTACAAACATCTTAAC	1739
QY	2157	caagaagaaacacagataatgtcacagaagaagatgtaacctgttctctgttatg	2216
Db	1740	CAAGAGGAACCGATGATGTCAAGACAGAGAAGAGCTATCTGTTTCTGTATTCCAT	1799
QY	2217	gagaacccaagctgtgtaacctcacagagaatalatacaagcttctcccaatccaagt	2276
Db	1800	GAGATCAAAAGCTGTGTAACCTGCGAGAGAAATATTCACGCTTCCTCCCAATCCG	1859
QY	2277	gtcgaacttgagatccaagatcccaagctccaacatagtcaacgaatcaatgtcat	2336
Db	1860	TTACAGCCCCAGGATCCAGAGTTCCAAAGCTTCTAACATATCAACAGCATCATAG	1919
QY	2337	gttttgaatgttgaggttgtaagttgtgttcagaaggtgcgaactgtaacttata	2396
Db	1920	GTTTTGTACTCTTGACGCTGTGCGTTTGTTCACAGAGTGGCATACTGGTACATTCT	1979
QY	2397	agcattggaacagacagactgtaactctcttcgctctctctctcgtgatatcactca	2456
Db	1980	AGTGTGGAGCAGACAGGAACTTCCTCCGCTCTCTCTGCGCTACACCTTCAACAC	2039
QY	2457	aaaatggtctatgaaagacaaactcaatcccatctcagagagaactgtctcatg	2516
Db	2040	AAAAAGGCTATGTGAAGACACTCACCTTCCTCCCTCTCAGGAAACAGGCTTTCTAT	2099
QY	2517	tgcataggaaaacccaagttctatgattctcgtgggtgcacaacatccaagactctga	2576
Db	2100	TCAATGGAACCCAGGCTCTGTGGGTCTTAGGGTCCACAACTCAGACTTGGGAAACGA	2159
QY	2577	ggcaatgacccgtctactgaaagttctcagttgtgacaagaacacgtgtgatatca	2636
Db	2160	GGGATGACAGCCTTACTGAAGGTGTATAGTTGTGACAGGAGCAATTGATATTATG	2219
QY	2637	gacagttatgaagatattcaagcactctgcgtgagtaaaacaatgaccttgaacca	2696
Db	2220	AACACTTATGTGAAGATATTCCAGGCTTCTTCTGAGTGAAGAAATGTCATTGAAC	2279
QY	2697	agctctcccaagatctcaagcacccctcagacatgcaaaagaattatgcaaccca	2756
Db	2280	-----	2779
QY	2757	coagttctgaaagccatcaacggaataaactcgttactactcttcagtcagatcaag	2816
Db	2280	-----GACATTAAGCTTCTACTTTTACAGCCGGAGGAAGAAC	2315
QY	2817	gaaatgtcatalgtatcacatatacagttgaaatgaagaagaaatttgaacattat	2876



Db 2316 AAAATGACTATGATGATATCTTCAACTGAACGAAGGAGAGAAGATTGTGACATTAC 2375  
QY 2877 gatagatgataaatacagagcccccgaagcgttcaaaaagaacagacatatattatc 2936  
Db 2376 GGTAGAGATGAAATCGAGAGCCCTCGCAGCTTTCAGAAAGAACCCGACCTATTTCAT 2435  
QY 2937 gctcaatgagagagcctctggtatataatgagaatgagtagctccacatgctctaagaac 2996  
Db 2436 GCTCGGGGAGAGAGCTCTGGGATTACGGGATGAGCCGATCCCGCGCGCTTAAGAAC 2435  
QY 2997 agggcctaagatggcagtgctccatgaatgagaagatgcttccagaattactgat 3056  
Db 2496 AGGGCTCAGAACGGAGAGAGTCTCGGTTCAGAAAGGAGTCTCCGGAATTTGCTGAC 2555  
QY 3057 ggcctcttactcagcccttataccgtgagaaactaaatgaacttgggaactctggg 3116  
Db 2556 GGCTCCTTCAGCAGCCGCTCTACCGGGGMACTCAACAAACCTTGGGCTTTGGGA 2615  
QY 3117 ccataataagacagaagttgaagataataatcatatgtaacttcaagaatcagcctc 3176  
Db 2616 CCTACATCAGAGCGGAGATTGAAGACATCATGTTCAAAACACGAGCGCT 2675  
QY 3177 cgtccatttccctctatctcagccttattctataggaagaatcagagcaagagca 3236  
Db 2676 CGTCCCTATTCTCTCTACGAGCCTTATTCTTATCCGGATGATCAGAGCAAGGGCA 2735  
QY 3237 gaacctgaaaaaaacttctcagagcctaatagaaacaaactactcttggaaatgca 3296  
Db 2736 GAACCTTGACACAACTTCCTCCAGCCAAATGAAACCAAGAACTTCTTGGAAAGTCCAG 2795  
QY 3297 catcataatgcaaccactaaagatgattgagcagcaagcctgggttattctcagat 3356  
Db 2796 CATACATGAGACCCACAGAAAGAGAGATTGACGCAAGCCTGGCTACTTTCTGAT 2855  
QY 3357 gctgaactgaaaaaagatgctgactcagagcctgattgagccctctgctctgcaact 3416  
Db 2856 GTTACCTGGAAAAAGATGTCACATCAGGCTTGATCGCCCTTCGATCTGCGCGCC 2915  
QY 3417 aacacactgaaacctgctctgagagaagaatgagcagataaggaattgctctgttttc 3476  
Db 2916 AACGCTGAAACGCTGCTCAGGTAACAAAGTACCGCTGCAAGAAATTCCTGTTTTTC 2975  
QY 3477 accactctgatgagacaaagctggtacttcaactgtaaaatagaaagaacactgag 3536  
Db 2976 ACTATTTTGTATGAGAAAGAGCTGTACTCTCACTGAATAATGGAAGAACTGCCGG 3035  
QY 3537 gctccctgcaatatacagatggaagatcccaactttaaagaaatatacgttccatgca 3596  
Db 3036 GCCCCTGCAACCTGAGATGAGAGCCCACTCTGAAGAAACATTCGCTTCATGCA 3095  
QY 3597 ataagtgtcaataatgatacactacactggtctagtaatgctcagatcaagaatc 3656  
Db 3096 ATCATATGGCTATGATGATGATACCTCCGCTTGTATATGGCTTCGAATCAAAAGATC 3155  
QY 3657 cgaatgatactgctcagcagcagcagcaatgaaacatccattcatcttcaatcagaga 3716  
Db 3156 CGATGGATCTGCTCAGACATGGCGAGCAATGAATAATTCATTTCGATTCTTTAGCGGA 3215  
QY 3717 catgtgtcactgtaaaaaaaagagatataaaatggcactgtatacaatctctatcca 3776  
Db 3216 CAGGTGTCAGTACGAGAAAGAGAGAGATATTAATGCGCGGTACATCTCTATCCG 3275  
QY 3777 ggtgttttgaagcagtggaatgttaccatccaagcttggaattggcggyggaatgc 3836  
Db 3276 GGTGTCTTTGAGACAGTGAATCTACCTCCAAAGTTGGAATTTGGCAATAAGATGC 3335  
QY 3837 cttaatggcagcactacatgctcgtgatgagcacactttctcgtgtatacgaataag 3896  
Db 3336 CTGATTTGGCAGACCTCGCAAGCTGGGATGAGCAGACTTCTCGGTGATACAGCAAGAG 3395  
QY 3897 tctgaagctccctgggaatgctcttgacaacattagagatttcaagatlaacagttca 3956  
|||||

Db 3396 TGTGAGCTCCACTGGAATGCTTCTGAGACATTAAGATTTTCAGATCAGCTTCA 3455  
QY 3957 ggaacaatatgacagtggtgcccacaagctggccagacttatttccgatacaaat 4016  
Db 3456 GGACGATATGAGAGTGGGCCCAAGCTGGCAGACTTCATATTTTCGGATCAATCAAT 3515  
QY 4017 gcttggagcaacagaagcccttctctggaatcaagtgatgctgttggccaatgatt 4076  
Db 3516 GCTTGAGACACCAAGATCCCACTCCTGATTCAGAGTGTCTGTTGGCACAATGATC 3575  
QY 4077 attcaggaatcaagaccgaagtgccogtcaagaagtttccagccttcaatctcag 4136  
Db 3576 ATTCCAGGATCATGACCAAGGATGCCCGTCAGAAATTTTCCAGCTCATCTCCAG 3635  
QY 4137 ttatcatatgataatgcttctgataaggaagaatggaagacttctgaagaattccact 4196  
Db 3636 TTTATCATCATGTACAGCTTTGAGAGGAACTGGCAGAGATTCACAGGGAATTCACG 3695  
QY 4197 gaaaccttaatgctctctcttggcaatgtagtcatctgagataaacaatatattt 4256  
Db 3696 GGCACCTTAATGCTCTTCTTTGGCAATGTGACCCATCTGGATTAACACAAATATTTT 3755  
QY 4257 aaccttcaattatgtctcgaatatacatcogtltgacccaactcataatgactcagc 4316  
Db 3756 AACCTCCGATTTGGTCCGCTACATCCGTTTGCACCAACATTTACAGATCCGACG 3815  
QY 4317 actcttcgcatggaatgtagtggtctgtatataatgttcgagcagatgccaatg 4376  
Db 3816 ACTCTTGATGAGTGTGATGGCTGTGATTTAAACAGTTTGCACCAATGCCCTGGGATG 3875  
QY 4377 gagaataagcaatatagatgacagattactcttcatctcattcaaatatgatt 4436  
Db 3876 CAGAAATAAGCATATGAGATCAGATCAGAGCCTCTCCACCTAAGCAATATATTT 3935  
QY 4437 gcaacctgtctccttcaaaaagctcgacttcaacctcaaggaagaaatgctcggaga 4496  
Db 3936 GCCACCTGTCTCCTTCACAAGCCCGACTTCACCTCCAGGGGCGAGAAATGCTGGCGA 3995  
QY 4497 cctcaggtgaataatccaaaagatggctgcaatggaacttccagaagaataaataagtc 4556  
Db 3996 CCCCAGTGAACAGCAGAGAGAGAGTGTGAGGTGAACCTCGCAAGAAACGCTAAAGTTC 4055  
QY 4557 acaggaataacttactcaggaagataaactctctgtcttcaacagcatgataatgaaagatc 4616  
Db 4056 ACAGGATACACACACCGAGCGCTGAATCCCTGCTCAGACAGATGATGCAAGAGATTTC 4115  
QY 4617 ctcaatccagcagcagaatgagcatalcagtggaactcttcttccagaatgycaaagta 4676  
Db 4116 CTGCTGTCAGTATGTCAGAGCGCGCGCTGACCTCTTCTTCAGAGCGGCCACAG 4175  
QY 4677 aaggttttcaaggaataatcaagactccttcaacactgttgtgaactctctagaccacg 4736  
Db 4176 AAGGTTTTTCAGGCAATCAGGATCTCTCCACCCCGTGTGAACGCTGTGAGACCCCGC 4235  
QY 4737 ttaatgctcgctacacttcaatcccccagaatgggttgagcagaatgctcagag 4796  
Db 4236 CTGTTCACGGGCTACCTGAGAGATCCACCACAGAGCTGGGCGCAGCATGCGCCTAGG 4295  
QY 4797 atggaagttctggctgagagcaagacctactgta 4835  
Db 4296 CTCGAGGTTCTAGATGTGAGCAGCAGAGATCTCTACTGA 4334  
|||||

RESULT 12  
US-09-037-601-38  
; Sequence 38, Application US/09037601  
; Patent No. 6180371  
; GENERAL INFORMATION:  
; APPLICANT: Lollar, John S.  
; TITLE OF INVENTION: Hybrid Human/Animal Factor VIII  
; NUMBER OF SEQUENCES: 40  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.



STREET: 5370 Manhattan Circle Suite 201  
 CITY: Boulder  
 STATE: Colorado  
 COUNTRY: USA  
 ZIP: 80303  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/037,601  
 FILING DATE: 26-JUN-1996  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: WO PCT/US94/13200  
 FILING DATE: 15-NOV-1994  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/212,133  
 FILING DATE: 11-MAR-1994  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/864,004  
 FILING DATE: 07-APR-1992  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Feider, Donna M.  
 REGISTRATION NUMBER: 33,878  
 REFERENCE/DOCKET NUMBER: 75-95F  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 303/499-8080  
 TELEFAX: 303/499-8089  
 INFORMATION FOR SEQ ID NO: 38:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 4334 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: not relevant  
 MOLECULE TYPE: cDNA to mRNA  
 HYPOTHETICAL: NO  
 ORIGINAL SOURCE:  
 INDIVIDUAL ISOLATE: Factor VIII lacking B domain  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 3..4334  
 US-09-037-601-38

Query Match	40.2%	Score 3189.8	DB 4	Length 4334
Best Local Similarity	83.5%	Pred. No. 0		
Matches 3692	Conservative	0	Mismatches 637	Indels 90
			Gaps	3
QY 420 atgcaaatgagctctccacccctgctctcttcctgctgctttagagatcttcgattagtc				
Db 3 atgacgactgagactctccacactgctgcttcttctgctgcttcttgccactcggctttagtc				
QY 480 accagaagaatactacctggtgctcagtgagactgtcataggagactatatagtcaaaagtga				
Db 63 atcagagagatctactactctggcgccagtggaactgtccttgagactaccggccaagtgaac				
QY 537 ctcggtgagctgcctcgtgtggaagcaaatcttcctctagagtgccaaatcctttccattc				
Db 123 ctccctgagactctcagcttgacacacagattttcctgtcttaccagccgacagactcttccgctg				
QY 597 aacacctgaagtgtgttaacaaaagaagactcgtttgtgtagaatataaggtttaccctttcaac				
Db 183 ggcctcgtacgtctctgacaaaagactgtgttctgtatgaatttcacgattcaactttttcac				
QY 657 atcgttaagcccaagagccacacctgagatcgttctgtcagtgccctacacatcccaagcttgagt				
Db 243 gttggcagggccacagggcacacatgagatgggtctgctgggtcttaccattccagggctgatt				
QY 717 tatgatatacagtggtcattatacacttaagaacatgagcttccactcctgtctcagcttccatgct				
Db 303 ttagcacacacgctgctgtttacccctgaagAACataggcttcttcaatccctgtatgtttcACgct				

[illegible]



QY	1857	ctgttgatataatattaagaatcaagcaagacacalataacatbaecttaecgaatc	1916
Db	1440	ctttttgattatattttaaagaattaaagccagccgacctatataacatctaccctcattgaaatc	1499
QY	1917	actgatgcgcctctctgtgtatcttaagaagatctccaaagcgtgtaaacatttgaagatc	1976
Db	1500	actgatgtcacccgcttttgccacccagggagacttttaaaagggttggaacatttgaaagac	1559
QY	1977	ttcccaattccgcagaagaaataatcoaaataaaatggacagtgtaactgtagaatctgg	2036
Db	1560	atggcaatttgcacagagagacgttttcagatnaatattgacagctgactgttgaaatctgg	1619
QY	2037	ccaactaaatcaagatcctcgggtgcctcgaccgctatlaactagtttgtaataatvgag	2066
Db	1620	ccaaaccaagttccgattccctcgctccgacccctgactactccagctccattaatctagag	1679
QY	2097	agaatctagcttcaagaccatcttgccctctctcatctgttacaagaatctgtgat	2156
Db	1680	aaagctctgcttggagactcattggccctctctctatctgctaacaaagattctgtaac	1739
QY	2157	caaaaggaagaaacagataaagtctacagaagaagaaatgcatctgttttcgtatattgat	2216
Db	1740	caaaaggaagaaacagatgactgacagacagaaagaaagctatctgttttctgtattgcgt	1799
QY	2217	gagaacogaagctgtgtaacctcacaagaaatatacaagcttctctcccaatccagctgta	2276
Db	1800	gagatcatcaagctgtgactgctccgacagaaattatcacagcttctctcccaatccgagatga	1859
QY	2277	gtgcaagcttgaagatccagaagcttccaagctccaacatcatgcaagaatcaatgtgctat	2336
Db	1860	tttcaagcccccagagatccagagattccaaagcttttcaatcatatcacagatcaatggctat	1919
QY	2337	gtttctatagtttgcagttgtcagtttgttggcaatgagtgagcaactgtgtatattctta	2396
Db	1920	gttttttgattagcttgcagcgtctgctgtttgtttggcaagagtgacactggttaattctta	1979
QY	2397	agcattvgagcaacagactgtaactccttctgtcttcttctctctgtatataacttaaacac	2456
Db	1980	agtttttgagacagacagacgactctctctccgcttcttctctgctgtacaccttcaaacac	2039
QY	2457	aaatgtctatataagaacacacacccatctccatcttccaagagaaactgcttctatg	2516
Db	2040	aaatattgcttattgaagacacacactccacccctgttcccttctcaaggaacacgctttcatg	2099
QY	2517	tcgaatggaacacccagctctatctgactctcgtgggtgcacaacatcaagacttccgaaacaga	2576
Db	2100	tcattatgaaaacccagactctctctgtgcttcttgaaggtgtccacactcagacttccgaacaga	2159
QY	2577	ggcatgacgccttactgtaaggttctctagttgtgtgaacaagaactggtatataacgag	2636
Db	2160	gggatgacacaccccttactgaaaggtgtattgttgacagaggaacattgctgattattatgac	2219
QY	2637	gacagttatagaatatatttcagatactctgtcgtgagtaaaacaatgcatatgaaccaaga	2696
Db	2220	aacacttatataaataattttccagagcttcttgcgtgagatgtgaagaaatgctattgaaccaga	2279
QY	2697	agcttccccaagattccaagacaacctagcactagagcaaaagcaatttaatgcaaccoca	2756
Db	2280	-----	2279
QY	2757	ccagctcttgaacgcacatcaacagggaaataactctgtaactcttcagtcagatcaagag	2816
Db	2280	-----gacattaaacctttcttcaagccggagaaagac	2315
QY	2817	gaaattgactatgataccatataagtttgaatgaaagaaagaaatattgacattat	2876
Db	2316	aaatttgactatgattatattttcttctcaactgaagaaagaaagaaatatttgacattttac	2375
QY	2877	gataagataaaatcaagacccccgcagcttcaaaagaaagaaagaaactattttat	2936
Db	2376	gctgagagatataaaatagacacccctgcagacttttcaagaaagaaacccgacactatttttatt	2435
QY	2937	gctgacagttgagaagcctctcgtgattatggaatgagtaagtaagctccacatglttctaaagaa	2996

Db	2436	GCCTGCGTGGAGCAGACTCTGGATTAACGGATGAGCGAATCCCCCGGCGCTAGAGAAAC	2495
Qy	2997	agggttcgaagtgtgcagtgtgtccctcagtttcaagaagtgtgtttccagaatttcagat	3055
Db	2496	AGGGCTTCAGAAACGGAGAGGTGCTCCGTGGTTTAAAGAGGTGCTTCCGGAAATTTTCTGCAC	2555
Qy	3057	gacctccttactcaagcccttaaccgtyggaagactaaatgaacatttggaactccctggg	3116
Db	2556	GGCTCCTTCACGGACCGCTGTTACCGGGGGAACTCAACAACACTTGGGGCTCTTGGGA	2615
Qy	3117	ccataataagaagagaagaatttgaagaataatataatgttactttcagaataatcaggccctc	3176
Db	2616	CCCTCACTCAGAGCGGAAGTTGAAGACACATCATGTGTAATCTTCAAAAACACAGCGCT	2675
Qy	3177	ggtccctcattccctctattcttagccttatttcttatgagaagaatcagaagcagaagca	3233
Db	2676	GGTCCCTAATTTCTTCTACTCGAGCGTTATTTCTTAATCCGATGATCAGGAGCAAGGGGCA	2735
Qy	3237	gaacctagaanaaactctgtcaagccctaaatgaacaacaaacttacttttggaaagtga	3296
Db	2736	GAACTGCACACAACTTCGTCCAGCCAAATGAACAAGACTTACTTTTGGAAAGTGACG	2795
Qy	3297	catcataatgcaaccacctaagaatgtgaatttgaactgcgaagccctggacttattcctgat	3356
Db	2796	CATCAGATGGCACCACAGAAAGAGATTTGACTGCAAAAGCGGGCGTACTTTCTGAT	2855
Qy	3357	gttgaactgtgaaaaagtgtgcactcgaagccgtaattggaaccctctgtgtctgcacac	3416
Db	2856	GTTGACCTGGAAAAAGATGTGCATCGAGGCTTATGCGGCCCTTCTGTATCTGCGCGCC	2915
Qy	3417	aacacactgaacccctgtctcatgtggagaagaatgaagaatcagaatattgcttgttcttc	3476
Db	2916	AACACCTGAAACGCTGCTACCGTAGACAGATGACCGTGCAGAAATTTGCTGTGTTTTTC	2975
Qy	3477	accactcttgaatgagaccaaagaactgttacttaaccgaaataatagtgaagaagaactgcag	3536
Db	2976	ACTATTTTGTGATGAGACAAAGAGCTGTACTTCACTGAAATGTGGAAAGAAAGACTGCCG	3035
Qy	3537	gtctccctgcgaataatccagatgtgaagaatcccaacttttaagaagaatatcgtctcattga	3596
Db	3036	GGCCCCGTGCCACCTGGAGATGGAGGACCCACACTGTAAAGAAACATTCGCTTCCATCA	3095
Qy	3597	atcaatgtcacaataatgagatatacaactcgggttgatgaatggcccaagatcaagaagatt	3656
Db	3096	ATCAATGGCTATGTGATGATTAACACTCCTGGCTTATGTATGTGCTACAGATTAAGAAGATC	3155
Qy	3657	cgatgtatctctgctcagcagtyggcagcaaatgaaataatcatctatcaattcaattgata	3716
Db	3156	CGATGTATCTGCTCAACCATGGGACAGACANTGAATAATATCATTCAGTTTCAATTTAGCGGA	3215
Qy	3717	catgtgttccactgaagaaaaaagaagatataaatggaacgtgcataatctcatcca	3776
Db	3216	CACCTGTTCAGTGTACGGAAAAAGAGAGATTAATAAATGGCCGTATCAATCTATTCGG	3275
Qy	3777	gggttttttgaagaagtggaatgttatcatctcaaaagcgtgaatttggcgggtgaaatgc	3836
Db	3276	GGGTCTTTTGAAGAGGAAATGCTAACCGTCCAAAGTTGGAAATTTGGCGATTAACAATGC	3335
Qy	3837	cttatgtgcagacatactaatctgtctggaatgaacacacttttctgtgtlaagaagaataag	3896
Db	3336	CTGATTTGGCGAGCACCTGGCAAGCTGGGATGAGACAGCACTTTCCTGTGTACACGACAGAG	3395
Qy	3897	ttgtcagaactcccttggaatgtgctctctgtgacacatlagagaatttcagattacagcttca	3956
Db	3396	TGTACAGCTCTCACCTGGGAATAGCTTCTGTGAGCAGCATTAGAGATTTTCAATCAACAGCTTCA	3455
Qy	3957	ggacaataatggaatgaggccccaagcctgycgcagacttcatattccggatcaatcaat	4016
Db	3456	GGACAGTATGACAGTGGGGCCCCAAAGCTGGCCAGACTTCAATTTATTCGGATCAATCAAT	3515
Qy	4017	ggcttgagacccaagagccctttctttgatccaagtgagatctgttgaaaccaatgatt	4076



```

Db 3516 GCGTGGAGCACCAGATCCCATCTCTGGATCAAGTGGATCTGTTGGCACCACCATGATC 3575
Qy 4077 attcaagcacaagaaccgaaggtgcgcctcagaagttctccagcctcacaatctccag 4136
Db 3576 ATTCAAGCACAATGACCAAGGTCGCCGTCAAGATTTCAGAGCTCTACATCTCTCCAG 3635
Qy 4137 ttatcatcatgatatagctctgaatgggaagaatggcagaacttatccaggaatccact 4196
Db 3636 TTTATCATCATGATAGCTCTGACGCGAGGAGAACTGGCAGAGTTACCGAAGGATTCACAG 3695
Qy 4197 ggaacctaagtgctcttttggcaatgtaattcctcgggataaacaatltttt 4256
Db 3696 GGCACCTTAATGATCTCTTTGGCATGTGGACGATCTGGGATTAACACATATTATT 3755
Qy 4257 aaccctcaatatgtctcgaatatacctgttgcacccaactcatalatagctcgcagc 4316
Db 3756 AACCTCGATTGTGGCTCGGATCATCGCTTGCACCCACACATTACAGCATTCGCACAC 3815
Qy 4317 actcttcgcatgagtgatgggtgtgtatataatagttgcagcatgcatgggaatg 4376
Db 3816 ACTCTTCGATGAGATTGATGGGCTGTGATTTAAACAGTTGCAGCATGCCCTGGGAATG 3875
Qy 4377 gagaagtaagaataatcagaatgacagatcttctcactcactcacttataatgttt 4436
Db 3876 CAGATTAAGCGATATCAACTCACAGATCACGGCTCTCTCCACCTAAGCATATATT 3935
Qy 4437 gccacctgtctccttcaaaaagctcgaactcactcaccagaaggagagtaatgctcgtaga 4496
Db 3936 GCCACCTGTCTCCTTCAACAAAGCCCGACTTCACTCCAGGGCGGACGAAATGCCCTGGCA 3995
Qy 4497 cctcagtgtaatatcaaaaagatgtgtcgaagtgtgacattccagaagaacaatgaatgc 4556
Db 3996 CCCCCGGGACAGCGAGAGAGTGGCTGCGAGTGGACCTCACAAGAGCGGTGAAGGTC 4055
Qy 4557 acaaggaatcactactcagaaggatataatctctgtctaccagatatgtatgtgaagtgc 4616
Db 4056 ACAGAGCATACACCCAGGAGGTGAAGTCCCTGCTACAGAGATGATGTGAAGGATTC 4115
Qy 4617 ctcaatccagcagtcagaatgagcagtagtgaactctcttcttccagaatggcaaatga 4676
Db 4116 CTCTGTCAAGTATGACAGAGGCGCGCGCTGAGCCCTGTTCTTCAGAGCGCCACAGC 4175
Qy 4677 aaggttttcagggaaatcaagaactccttcacacctgtgtgaactctcttagaccacgcg 4736
Db 4176 AAGGTTTTTCAGGGCAATCAGAGACTCTTCACCCCGTGTGTAAGCTCTGACCCCGC 4235
Qy 4737 ttatgactcgtactaccttcgaattcaaccccaagatgttggtcacagaatgctcctgaag 4796
Db 4236 CTGTTCAAGCCTTACCTAGAGATCCACCCACGAGACTGGGCGCCAGCACATCGCCTGAGG 4295
Qy 4797 atggaagttctggtgctgcgaagcaagaagactctactga 4835
Db 4296 CTCGAGGTTTAGATGTGAGGACACAGATCTCTACTGA 4334

```

```

RESULT 13
US-08-893-327-15
; Sequence 15, Application US/08893327
; Patent No. 6020192
; GENERAL INFORMATION:
; APPLICANT: Zoletukhin, Sergei
; APPLICANT: Hauswirth, William W.
; APPLICANT: Muzyczka, Nicholas
; TITLE OF INVENTION: Humanized Green Fluorescent Protein
; TITLE OF INVENTION: Genes and Methods
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P. O. Box 4433
; CITY: Houston
; STATE: TX
; COUNTRY: USA
; ZIP: 77210-4433

```

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/893.327
; FILING DATE:
; CLASSIFICATION: 514
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/588,201
; FILING DATE: 18-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Kitchell, Barbara S.
; REGISTRATION NUMBER: 33,928
; REFERENCE/DOCKET NUMBER: UFLA:062\KIT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (713) 789-2679
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6253 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 988..1701
; US-08-893-327-15

```

```

Query Match 37.6%; Score 2985.2; DB 3; Length 6253;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 3000; Conservative 0; Mismatches 3; Indels 3; Gaps 1;

Qy 4916 agatctagaagacccttagtgaatgttggcactcctctctgogctcgtcgtca 4975
Db 3250 AGATCTAGAACCCCTAGTAGAGTGGGACACTCCCTCTGCGGCTGCTGCTCA 3309
Qy 4976 ctgagagcccgccggaagaagcccggtcgtggaacttgtgtccgcgcctcaatga 5035
Db 3310 CTGAGGCGCCCGGCAAAAGCCCGGCGCTCGGGCACCCTTTGGTGGCCGCTCATGTA 3359
Qy 5036 gcaagcgaagcgcaagaaggaagtgccaaacccccccccccccccctcgaaccagc 5095
Db 3370 GCGAGCGAGCGCGAGAGAGGAGTGGCCAAACCCCCCCCCCCCCCTCAGCC---C 3426
Qy 5096 tgcattaatgaatcgccaaagcgcggtggaagagcggtttgctatgtgggtcttcgcg 5155
Db 3427 TGCATTAAATGAATCGGCCAACGCGCGGGAGAGGGGTTTCGTAATGGCGCTCTTCG 3486
Qy 5156 ctctcgtcactaacctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 5215
Db 3487 CTTCCTCGCTACCTACTACCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCG 3546
Qy 5216 actcaagcggtlaatcagttalccacagaatcaagggaataacgaggaagaacatgt 5275
Db 3547 ACTCAAGCGGATATAGGTTATTCACAGATCAGAGGATTAACGAGAAAGAAATGT 3606
Qy 5276 gagcaaaagcgcaagaagcgcaagaacgttaaaagcggttctgtcgttcttc 5335
Db 3607 GAGCAAAAGGCGCAGCAAAAGGCCAGAACCGTAAAGGCGCGCTTCGCTGCGTTTTC 3666
Qy 5336 ataggctcgcgcgccttcagagacatcaaaaatcagcgtctcaagtcaagaagtgagcga 5395
Db 3667 ATAGGCTCGCCGCCCTTACAGCATTCACAAAATTCAGCTCAGTACAGAGTGGCGGAA 3726
Qy 5396 accgcagagactaataagataaccagagcttccctcgtgaagctcctcgtcgtcctc 5455
Db 3727 ACCGACAGAGATTAAGATATACAGAGCGCTTCCCGCTGGAAGCTCCCTGCTGCTCTC 3786
Qy 5456 ctgttcgaaccctcgcgtctacggataactcgtcgcgtctctcctcgtgggaagcgtg 5515

```



```
Dh 3787 CTGTCGACCCCTCCGCTTACCGGATACGTCGCCCTTCTCCCTTCGGGAACGCTG 3846
Qy 5516 cgtcttcaatgctcagctcgttagtatactcagtcgttagtgcgttcgccaagc 5575
Dh 3847 CGCTTTCATGCTCACGCTGTAGTATCTCACTTGGTGTAGTGGCTTGGCTCCAGC 3906
Qy 5576 tggcgtgtgcaagAACCCCCCTCAGCCGACCGGTGGCTTATCCGTAATCATC 3966
Dh 3907 TGGGCTGTGTCAGAACCCCCCTCAGCCGACCGGTGGCTTATCCGTAATCATC 3966
Qy 5636 gctcttagtccaaccggttaagacagactatcgccaactggaagccaactgtaaca 5655
Dh 3967 GTCTTGAGTCCACCCGTAAGACACGACTATCGCACCTGGCAGCCACTGCTAACA 4026
Qy 5696 ggaatacagaagaggtatgtagcggtgctacagatctttaaattggtgactact 5755
Dh 4027 GGATTAGAGGCGAGGATGTAGCGGTGTACAGAGTTCTTAAGGTGGGCTTAATCT 4086
Qy 5756 acggtacaactagaagagcaatattggtatctgcgtctgtaagccaactgctcg 5815
Dh 4087 ACGGCTACACTAGAGAGACAGTATTGGTATCTCGCTCTGCTGAAGCACTTACCTTG 4146
Qy 5816 gaaagaagttgtagctcttgatcgcgcaaaacaaccgcgtgtagcggtgttctt 5875
Dh 4147 GAAAAAGATTGAGCTCTTGATCCGGCAAAACACCGCTGTAGCGGTGTTT 4206
Qy 5876 ttofttcaagcagaatagtcgcgagaanaaagagatctcaagaatcttgatctc 5935
Dh 4207 TTGTTTCAGACGACAGATTAAGCGCAGAAAAAGATCCAGAAAGATCTTTGATCT 4266
Qy 5936 ttctaaagggctcgaagctcgaatggaacgaactcagcttaaggatttgctaga 5995
Dh 4267 TTTCTACGGGGTCTGAGCGTCAGTGAACGAAACACGCTTAAGGATTTTGGTCATGA 4336
Qy 5996 gattatacaaaagatctcactagatcctttaaattaaatgaagtttaactaa 6055
Dh 4327 GATTATCAAAAAAGATCTTCACTAGATCCCTTTAAATTAATAATGAAGTTTAAATCA 4386
Qy 6056 tctaaatataatgagtaactggtctgacagttacaaatgcttaactgaaggcac 6115
Dh 4387 TCTAAATATATATGATTAACCTTGTGTGACAGTTACCAATGCTTATATCAGTAGGCAC 4446
Qy 6116 ctactcagcagatctgctatcttcgtcatcatcagttgctgactcccgctgctaga 6175
Dh 4447 CTATCTCAGGAGATCTGCTATTTGCTTCATCCATAGTTGGCTGACTCCCGTCCGTAGA 4506
Qy 6176 taactacgatacgggagggcttaccatctggcccagtgctgcaatgataccgagacc 6235
Dh 4507 TTACTAGATACGGGAGGGCTTACCATCTGGCCCCAGTGTGCAATGATACCGCGAGACC 4566
Qy 6236 caggtacaacgggtccagatttatacgaataaaccagccagccgggaaggccgaagcca 6295
Dh 4567 CACGCTACCGGGCTCCAGATTATACGAATAAACACGACGCGGGAAGGGCCAGCGCA 4636
Qy 6296 gaattggtcctcgaacttalcgcgtccatcagctcatatattgcttgcgggaagcta 6355
Dh 4627 GAATGTGCTGCAACTTTATCCGCTCCATCCAGTCTATTAATTTGTTGCCGGGGAAGCTA 4686
Qy 6356 gagttaagtagtcgcagttaaatagtttgcgcaagcttgctgacattgctacagcatcg 6415
Dh 4687 GAGTAATAGTATCCCGAGTTAATAGTTGGCAACGTTGTTGCCATTTGCTACAGCATACG 4746
Qy 6416 tgggtcaacgctgctgttggtagtgcgttcacatcagctccgggttcccaagatcaagggc 6475
Dh 4747 TGGTGTACGCTGCTGCTTGGTGTAGGCTTCATTCAGCTCCGGTTCCCAAGATCAAGGC 4806
Qy 6476 gagttaacatgcccccatgctgtgcaaaaaaagcgttagctcccttcgctccgactcg 6535
Dh 4807 GAGTTACATGATCCCGCATGTTGTGCAAAAAAGCGTTAGTCTCTTGGTCTCCGATCG 4866
Qy 6536 ttgtcagaagtaagttggccgagtgctatcaactcagtgtatggaagcactgataatt 6595
Dh 4867 TTGTCAAGATTAAGTTGGCGCATGTTATCATCATGATGTTATGGCAGCACTCATTAAT 4926
Qy 6596 ccttactatcatgcatcgcgtaagatgcttcttgtagactggtgagttactcaaccaagt 6655
Dh 4927 CTCTTACTGTATATCCATCCGTAAGATGCTTTTCTGTGACTGGGTAGTACTCAACCAAGT 4986
Qy 6656 catctgagaatagtgtagtcgagcagagttgctcttgcggcgctlcaatacgggata 6715
Dh 4987 CATCTGGAATAGTATGATGGCGACCGAGTGTCTGTGGCCGGCGTCATATACGGGATA 5046
Qy 6716 ataccggccacatagcaagaactttaaagtgtccatcatctggaaaacgttctcggggc 6775
Dh 5047 ATACGGGCGCCATAGCAAGAACTTTAAAGTGCATCATCTGGAAGAAAGTTCTTCGGGGC 5106
Qy 6776 gaaacctcaagatcttaccgtgttgtagatccagttcagttgaatcccaactcgttcac 6835
Dh 5107 GAATACTCTCAAGATCTTACCGCTGTGTAAGATCCAGTTGATATACCACTGCTGAC 5166
Qy 6836 ccaactgatcttcagcatcttctacttcaaccaagcgttctcgggtgagtgacaanaacaggaa 6895
Dh 5167 CCAACTGATCTTGACATCTTTACTTTACACACGTTTCTGGGTGAGCAAAAAACAGGAA 5226
Qy 6896 ggcataatgccgcaaaaaaggaatgaaggcgacacggaatggttgaatactactctc 6955
Dh 5227 GGCATAATGCCGCAAAAAAGGAATAAGCGCACACGGAATGTTGAATACTCATACTCT 5286
Qy 6956 tccctttcaaatatttgaagcatcttaagggttatgtctatgtcgaatgaagcatatat 7015
Dh 5287 TCTTTTTCATATATTTAGACATTTATATCAGGTTTATGTCTATAGCGGATACATAT 5346
Qy 7016 ttgaatglatltaagaaataaacaatagaaggtctccgcgcacattcccccgaagaatgctc 7075
Dh 5347 TTGAATGATTTAGAAAAATTAACAATAAGGAGTTCCGCGCACATTTCCCGGAAGAAGTC 5406
Qy 7076 caactgaagttcaagaacacattatcatgacattaaactataaataatagcggtatca 7135
Dh 5407 CACCTGACGCTTAAGAAACCAATTTATCATGACATTAACCTAATAAATAGCGGTATCA 5466
Qy 7136 cgaagcccttgcgtcgcgcggtctcggtgaatgaacgttgaacactcgtacacatgacgc 7195
Dh 5467 CGAGGCCCTTTCGTCGCGGCTTCGGGTGATGACGGGTGAAGAAACCTGTGACACTGTGACGC 5266
Qy 7196 tcccggagacggtacacagctgtctgtaagcgaatgcccggagacagaacagcccgtagc 7255
Dh 5527 TCCCGGAGACGCTACAGACTTGTCTGAAGGAGATGCCGGAGACAGAACCCGCTCAGG 5586
Qy 7256 ggcgctcagcgggtgttgcggcgtgtcggggctggttaactatgcygcacatgaagcaga 7315
Dh 5587 GCGGCTCAGCGGCTGTTGGCGGGGTGTCGGGCTGCTTAATGATGCGCATCAGAGCAGA 5646
Qy 7316 ttgtactgagatgacacatatgagtggtgtaataacgcgacagaatggtgaagggaaat 7375
Dh 5647 TTGTACTGAGATGACCATATGCGGTGTAATAACGCGACAGATGGTAAAGGAAAT 5706
Qy 7376 accgcatacaggaatctgtaaacgtttaaatttctgtaaaatctcgctttaaatttctgta 7435
Dh 5707 ACCGCATCAGGAATTTGAAACGTTAATATTTGTTAAATTTGCGTTAAATTTTGTGTA 5766
Qy 7436 aatcagctcaatttcttaaccaatagccgaataatcggcataaataccctataataaaga 7495
Dh 5767 AATCAGCTCATTTTAAACCAATAGGCCGAATGCGCAAAATCCCTTATTAATCAAAAGA 5826
Qy 7496 atagacgagataaggtgtggtgtgttccagtttggaaacaaagatccaatataaaga 7555
Dh 5827 ATGAGCCGAGATAGGTTGAGTGTGTTCCAGTTTGGAAACAGAGTCCACTATTAAGAA 5886
Qy 7556 cgttgaactcaacgltcaaaagggcgaaaaacgltcatcaagggcgaatgcccactacgltga 7615
Dh 5887 CGTGACTCCAAAGCTCAAAAGGGCAAAACCGTTATACAGGGCATGCGCCACTACATGTA 5946
Qy 7616 accataacctaatcaagtttctggggtgcgaagtgccgtaaaagcaactaatcggaacc 7675
Dh 5947 ACCATCACCCCTTAATCAAGTTTGTGGGTGAGGTGCCGTAAGAACACTAATCGGAACCC 6006
```



QY 7676 taaaggagcccccgattagagcttgacgaggaaagccgagcaactgctgagagaaggga 7735  
| | | | |  
Db 6007 TAAAGGAGAGCCCCGATTGAGCTTTGACGCGGAAAGCCGCGGAACGTGGCAGAAAGGA 6066  
QY 7736 agggaaagaacgaaagagcgagcgctagggcgctgtgaaatgtgacgttaagcttgg 7795  
| | | | |  
Db 6067 AGGGAAAGAAAGCAAGAGAGCGGCGCTAGAGGCGCTGCGCAAGTGTAGCCGTACCGCTGGC 6126  
QY 7796 cgtacccacacacccgagcgcttaatgacgctacacagggcgagctcgccatctgcg 7855  
| | | | |  
Db 6127 CGTAACCAACCAACACCGCGCGCTTAATGCGCGCTACAGGCGCGGCTCGCGCATTCGCGC 6186  
QY 7856 attcaagctacgcaactgttgggaagggcgatcggtgcgacctctcgctattacgca 7915  
| | | | |  
Db 6187 ATTGAGGCTAGCAACTGTGGGAGAGGCGATCGGTGCGGCGCTCTTCGCTATTACGCCA 6246  
QY 7916 gctggc 7921  
| | |  
Db 6247 GCGTGC 6252

RESULT 14  
US-08-893-327-17  
; Sequence 17, Application US/08893327  
; Patent No. 6020192  
; GENERAL INFORMATION:  
; APPLICANT: Zolotukhin, Sergei  
; APPLICANT: Hauswirth, William W.  
; APPLICANT: Muzyczka, Nicholas  
; TITLE OF INVENTION: Humanized Green Fluorescent Protein  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Arnold, White & Durkee  
; STREET: P. O. Box 4433  
; CITY: Houston  
; STATE: TX  
; COUNTRY: USA  
; ZIP: 77210-4433  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentln Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/893,327  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/588,201  
; FILING DATE: 18-JAN-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kitchell, Barbara S.  
; REGISTRATION NUMBER: 33,928  
; REFERENCE/DOCKET NUMBER: DFLA:062\KIT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (512) 418-3000  
; TELEFAX: (713) 789-2679  
; INFORMATION FOR SEQ ID NO: 17:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 6280 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 988..1728  
; US-08-893-327-17

Query Match 37.6%; Score 2985.2; DB 3; Length 6280;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 3000; Conservative 0; Mismatches 3; Indels 3; Gaps 1;

QY 4916 agatctaggaacccttaatgtagtctggccactccctctctgcgctcgtctca 4975  
| | | | |  
Db 3277 AGATCTAGGAACCCCTAGTGAATGAGATTGGCCACTCCCTCTGTGGGCGCTGCTGCCTCA 3336  
QY 4976 ctgaagcgcccgagcaagccggcgctcgagcacttggctgcgcccgtcaatga 5035  
| | | | |  
Db 3337 CTGAGGCGCCCGGGCAAGGCCGGGCGCTCGGCGCACTTTGTGTGGCCCGGCTCACTGA 3396  
QY 5036 gcgagcgaacgagcagagagagtgagcacaacccccccccccccccctgcagccagc 5095  
| | | | |  
Db 3397 GCGAGCGAGCGCGCAGAGAGGAGTGGCCAAACCCCCCCCCCCCCCTTCACACC---C 3453  
QY 5096 tgcattaaatgaaatcgagcaacgagcgaggagagcggtttgcatttggcgctcttcg 5155  
| | | | |  
Db 3454 TGCATTAAATGAATCGGCCAACGCGGGGAGAGGGGTTTGCGTATTTGGGCGCTCTTCG 3513  
QY 5156 ctctcgtcactgactcgctgcgtgcgtcgttcgttcgctgcgagcggtatcagctc 5215  
| | | | |  
Db 3514 CTCTCTCGCTCACTGACTGCTGCGCTCGTCTGCTGCGCTCGGCGAGCGGTATCACTC 3573  
QY 5216 actcaaggcggttaataggttatccagaatacaggagataacgaggaagacatgt 5275  
| | | | |  
Db 3574 ACTCAAGCGGTATATGATTATCCACAGATCAGGCGATTAACGAGAAAGCAATGT 3633  
QY 5276 gaggcaaaagccaagaaagccaagaaacgtaaaagccgcttgcgtgctttcc 5335  
| | | | |  
Db 3634 GAGCAAAAGGCGCAGAAAGGCGAGAACCGTAAAAAGCGCGCTTCTGCTGCTTTTTC 3693  
QY 5336 ataggctccgccccctgacgagcatcaaaaaatcgacgtcgaatcgaagtgagcgaa 5395  
| | | | |  
Db 3694 ATAGGCTCGCGCCCCCTTACGAGCATCAAAAATCGACGCTCAAGTCAGAGTGGGGAA 3753  
QY 5396 accgacagagcatataaagataacagcgcttcccccggagagctcctctgtgccttc 5455  
| | | | |  
Db 3754 ACCGACAGAGATTATTAAGATACAGGCGTTTCCCTTGAGAGCTCCCTGCTGGCTCTC 3813  
QY 5456 ctgttcacgacctgcgcttaccgatactcgtccgcttctcccttcgggaaagctg 5515  
| | | | |  
Db 3814 CTGTTCCGACCGTCCGCGTTACCGGATACGTGTCGCGCTTCTCCTCGGGAAGCGTGG 3873  
QY 5516 cgcttctcaatgctcgaacgctgtaggtatcctcaagttcggtgtagtgctgcctcaagc 5575  
| | | | |  
Db 3874 CGCTTCTCAATGCTCAGCGCTGTAGTGTCTCAGTTCGTTAGTGTGTTGCTGCCACAGC 3933  
QY 5576 tgggctgtagtgcaagcaaccccccgctcagccgagccgctgagcttaccgtaactac 5635  
| | | | |  
Db 3934 TGGGCTGTGTGACAGCAACCCCGCTTCA6CCCGACCGCTGCGCCTTATCCGTTAATTC 3993  
QY 5636 gtcttgagtcacaaccggtlaagacacgactatcgccactgagcagccaactgttaca 5695  
| | | | |  
Db 3994 GTCCTGAGTCCAAACCCGGTAAAGACAGACTTATGCGCCACTGGACGACCACTGTAAACA 4053  
QY 5696 ggaatagcagagcgaagtatctaggcggtgctacaagatcttgaagtgtgtgcttaact 5755  
| | | | |  
Db 4054 GGATTAGCAGAGCGAGGTATGTAGCGCGTCTTAAGATGATTTTAAAGTGGGCCCTAACT 4113  
QY 5756 acggtctacagaaaggaacagatattgtgtactcgcgtcgtcctaagccagttactctg 5815  
| | | | |  
Db 4114 ACGGCTACACTAGAGAGAGAGATTTGTGTATGTGGCTCTGCTTAAGCCAACTTACCTTCG 4173  
QY 5816 gaaaaagattgtagctcttgatccggaacaaacacacccgctgtagcggtgtttt 5875  
| | | | |  
Db 4174 GAAAAAGAGTTGTTAGTCTTGTGATCCGCAAAACACACCGCTGTGTAGGGGTGTTTTT 4233  
QY 5876 ttgcttgcaagcagcagattacgagcaaaaaaagatctcaagaagatcccttgatctc 5935  
| | | | |  
Db 4234 TTGTTTGCAGACAGCAGATTACGGGCAGAAAAAAGGATCTCAAGAAAGATCTTTTGTATCT 4293  
QY 5936 ttctcaggggtctgaagctcagtgagaaacgaactcaagcttaaggagatttgtctatga 5995  
| | | | |  
Db 4294 TTTCTACGGGGTCTGACGCTCAGTGTGAACGAAACTCACGTTTAAGGATTTTGTCTATGA 4353



QY 5996 gattatcaaaaagatcttcacccatagatccctttaaataaaatgaagttttaaataa 6055  
|||||  
Db 4354 GATTATCAAAAAGATCTTCACCTAGATCCCTTTAAATTAATAATGAAGTTTAAATCA 4413  
QY 6056 tctaaagtataatgaagtaaaacttgctcgaagttacaaatgcttaacgaatgaagcac 6115  
|||||  
Db 4414 TCTAAAGTATATATGAGTAAACTGTGCTGACACTTACCAATGCTTAATCAGTAGAGGC 4473  
QY 6116 ctactcaagagatctgtctattcttcttcacatcaagtgttcgctgctcccgctgtaga 6175  
|||||  
Db 4474 CTATCTGACGATCTGTCTATTTGTTATCATCATAGTTGGCTGACTCCCTGCTGTAGA 4553  
QY 6176 taactacgatacgggaaggagcttaccatctggcccgagtgctcgaatgatacgcgagacc 6235  
|||||  
Db 4534 TAATCAGATACGGGAGGAGGCTTACCATCTGGCCCAAGTGCATATATACGGGAGAGCC 4593  
QY 6236 caagctacacggctccagattatcagcaataaaccagccagccggaaggcgagcgca 6295  
|||||  
Db 4594 CACGCTACCGGCTCCAGATTTATCAGCAATTAACCAAGCCAGCCGGAAGGCGCAGCGCA 4653  
QY 6296 gaagtgtcctcgcaacttaccgctccacatccagttctatatttctgttcgggaagcta 6355  
|||||  
Db 4654 GAATGTCCTCGACACTTTATCCGCTCCATCCAGTCTATTAAATGTTGCCGGAAGCTA 4713  
QY 6356 gaagtaagtagtcgcgacgttataagtttcgcgaacgttgctgcacattgctacagacatcg 6415  
|||||  
Db 4714 GAGTAAGTAGTTCGCCAGTTAATAGTTTGGCAACGTTGTTGCCATTGCTACAGCATTCG 4773  
QY 6416 tgtgtcagcgtcgtctgttgataggtcttcacatccagctccggttcccaacgatcaagc 6475  
|||||  
Db 4774 TGGGTGACGCTCCTGCTGTTGGTATGGCTTCATTCACGCTCGGTTCCCAACGATCAAGGC 4833  
QY 6476 gaattacataatcccccattgtatgcgcaaaaagaggttaagctcttcgttccttcagtcg 6535  
|||||  
Db 4834 GAGTTAATATGATCCCATAGTTGTCAGAAAAGGGGTTAGCTCTTGCTGGTCTCCGATCG 4893  
QY 6536 ttgcagaagtaagttgcgcgcaqtgatatacactacgtgtatagcagacactgacataat 6595  
|||||  
Db 4894 TTGTACAGAGTAAGTTGGCCGCACTGTTATCATCATGATGTTATGGCAGCACTGCATATT 4953  
QY 6596 ctcttaactgtcatatgcgtcgttaagaatgcttctgtgaactgtgtgaactcaacgaatg 6655  
|||||  
Db 4954 CTCTACTGTCATCCCATCCGTAAGATGCTTCTGTGACGTGGAGTACTCAACCAAGT 5013  
QY 6656 catctcagaagtagtgaatgcgagcgagcttgcttccgcgcgctcaataagagata 6715  
|||||  
Db 5014 CATCTGTAGAAATAGTATATGCGGACCGAGTTGCTCTTGCCGCGCATATAGGGATTA 5073  
QY 6716 ataccgcgcacacatagacagaactttaaagtgcatacatctgtaaaaaagcttctcgaggc 6775  
|||||  
Db 5074 ATACCGGGCCCATAGCAAGAACTTTAAAGTGCTCATCATTTGGAACGTTCTTCGGGGC 5133  
QY 6776 gaaacatctcaaggaatcttaccgctgttgagatccagttcgatgtaacccaactcgtgcac 6835  
|||||  
Db 5134 GAAACATCTCAAGAGATCTTACCGCTGTGAGATCCAGTGCATTAACCCACTGTTCCAC 5193  
QY 6836 ccaactgatctcgaacatcttcttacttaccacagcgttctcggttgaggcaaaaaaagaagaa 6895  
|||||  
Db 5194 CCAACTGATCTTCAGCATCTTTTACTTTACACAGCGTTTCTGGGTGAGCAAAAACAAGAA 5253  
QY 6896 ggcacaaatgcgcgcaaaaagagaaatgaaggcgacacggaatgttgatactactct 6955  
|||||  
Db 5254 GGCMAAATGCCGCAAAAAGGAATAGGGGAGACAGGAATGTTGAATCTCATCTCT 5313  
QY 6956 tccctttcaatatattgaagcaattatcaagggtatgtctcatabgagcggaatacat 7015  
|||||  
Db 5314 TCCATTCTTCAATATATTATGAAGCAATTATCAGGGTTATTGTCATATAGGGGATACATAT 5373  
QY 7016 ttgaattatttaaaaaataaacaataagaggttcgcgcacacttccccgaaaaatgac 7075  
|||||  
Db 5374 TTGAAATGATTATTAGAAAAATTAACAAATAGGGGTTCCGCGCACATTTTCCCGAAAAATGTC 5433  
QY 7076 caactgagcttaagaaacatatatcatgacatlaacctataaataatagcgatca 7135

Db 5434 CACCTGACGTAGAAACCATATATATCATGACATTAACTATATAAATAAGGCTATCA 5493  
|||||  
QY 7136 cgaggcccttcgtctcgcgcgcttgcgtgatgaacggtgaacacttgacatagcagc 7195  
|||||  
Db 5494 CGAGCCCTTTCGTCTGCGCGGTTCCGTGATGACGGGTGAACCACTTGACATGACAGC 5553  
QY 7196 tcccgagaagcggtacagactgtctgttaagaggttgcccggaaggaagaccccgtaag 7255  
|||||  
Db 5554 TCCGGAGACGCTACAGCTGTCTGTAAGCGGATGCGGGAGAGCAACGCCGTCAGG 5613  
QY 7256 ggcgctcagcggtgtctgcgcggttcggttcgaactacatgcgcacatcaagagaga 7315  
|||||  
Db 5614 GCGGTCAGCGGGGTGTGGCGGCTGTGGGGCTGCTTAATATAGCGGATCAAGCAGA 5673  
QY 7316 ttgactgagagtgacacatatgcggtgtgtaataaccgcacagatgcgtgaaggaat 7375  
|||||  
Db 5674 TTGTACTGAGAGTGCACCAATATGCGGTGTAATTAATGCGTTAAATTTTGTGA 5733  
QY 7376 accgcatcagaagaattgtaaacgttaatttgctttaaacttcgcgtttaaatttgta 7435  
|||||  
Db 5734 ACCCATCAGGAATTTTAACGTTAAATGTTAAATTTTGTAAATGCGTTAAATTTGTGA 5793  
QY 7436 aatcagctatlttlttaacacatagccgaatcgcgaataatccctataaatcaaaaaa 7495  
|||||  
Db 5794 AATCAGCTCATTTTAAACCAATAGCGCAAAATGCGCAAAATCCCTTAATTAATCAAAA 5853  
QY 7496 atagaccgaagataggttgatgtgtctcagtttggaacaagatgccaactataaaga 7555  
|||||  
Db 5854 ATAGCCGAGATAGGTTGATGATGTTGTTCCAGTTTGGACAAAGAGTCCACTATTAAAGAA 5913  
QY 7556 cgttgactccaagtcgaaggcgcaaaaacggtctatcagggagatggccactagtg 7615  
|||||  
Db 5914 CGTGACTTCAACGTCNAAGGGCGAABACCGTATACAGGGCATGCGCCACTACGTGA 5973  
QY 7616 accataccctaaatcaagttlttggttcgaagtgccgtaaaacacataatcggaaacc 7675  
|||||  
Db 5974 ACCATCAACCCCTAATCAAGTTTGTGGGTCGAGGTGCGATGAACACTTAATGGAACCC 6033  
QY 7676 taaaggagcccccgatlttagagcttgacgggggaagcggcggaagcttgcgagaaga 7735  
|||||  
Db 6034 TAAAGGAGGCCCCGATTTGAGCTTACGCGGGGAAACCCGCGAATCTGGCGAATAAGGA 6093  
QY 7736 agggaaagaagcggaaggagcgagcgcttagggcgcttgcaagtgtaacggltcagcgcg 7795  
|||||  
Db 6094 AGGGAATAAAGCCAAAGAGAGCGGCGCTTAGGGCCCTGCAAGTGTACGGTCAAGCTGCG 6153  
QY 7796 cgtlaaccacacacccgcgcgcgttlaatgcccgtctacaggcgcgctgcgcacatcgcc 7855  
|||||  
Db 6154 CGTAACCAACACACCCCGCGGCTTAATGCGCGCTTAAGGGCGCGTACAGGGCGCGCATTCGCGC 6213  
QY 7856 atcagagctacgcgaactglttggaaggcgatcggtgcgggctcttcgtataagcga 7915  
|||||  
Db 6214 ATTCAAGGCTACGCAACGTTGGGAAGGGCATCGTGCGGGCCCTTCGCTATTACGCCA 6273  
QY 7916 gctgcgc 7921  
|||||  
Db 6274 GGCTGC 6279

RESULT 15  
US-08-893-327-19  
; Sequence 19, Application US/08893327  
; Patent No. 6020192  
; GENERAL INFORMATION:  
; APPLICANT: Zoletukhin, Sergei  
; APPLICANT: Hauswitz, William W.  
; APPLICANT: Muzyczka, Nicholas  
; TITLE OF INVENTION: Humanized Green Fluorescent Protein  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Arnold, White & Durkee



STREET: P. O. Box 4433  
CITY: Houston  
STATE: TX  
COUNTRY: USA  
ZIP: 77210-4433  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/893,327  
FILING DATE:  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/588,201  
FILING DATE: 18-JAN-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Kitchell, Barbara S.  
REGISTRATION NUMBER: 33,928  
REFERENCE/DOCKET NUMBER: UFLA.062\KIT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (512) 418-3000  
TELEFAX: (713) 789-2679  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6280 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 988..1728  
US-08-893-327-19

Query Match 37.6%, Score 2985.2; DB 3; Length 6280;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 3000; Conservative 0; Mismatches 3; Indels 3; Gaps 1;

4916 agactctggaacccctctgtagtgagtggtgacacccctctctgagcgtcgtcgtca 4975  
|||||  
3277 AGATCTAGGAACCCCTAGTAGTAGGAGTTGGCCACTCCCTCTGCGCTGCTGCTCA 3336  
|||||  
4976 ctgagcgcccgccggaagcccgagcgctcggtcgagccttctgctccgagctca 5035  
|||||  
3337 CTGAGCGCCGCCGCGCAAGCCGCGGCGCTCGGCGCACTTTGGTGGCCGCGCTCAGTGA 3396  
|||||  
5036 gcgagcgagcgcgcaagagagagtgccaaacccccccccccccctcagccagc 5095  
|||||  
3397 GCGAGCGAGCGCGCAAGAGGAGTGGCCAAACCCCCCCCCCCCCCTCGAGGCC--C 3453  
|||||  
5096 tgccttaatgatacgccaaacgcgagggagagagcggttctgatacttgagccttcgc 5155  
|||||  
3454 TGCATTATGATGCGCCAAAGCGCGGAGAGCGGTTGGATTTGGGCGCTCTTCG 3513  
|||||  
5156 ctctctctactactgactcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 5215  
|||||  
3514 CTTCCTGCTGCTGACTGACTCGTGGCGCTGGTGGTTCGCTGGCGGAGCGGTATCAGTCA 3573  
|||||  
5216 actcaaaagcggtatacagttatccacagaatacagggagataacgaggaagaacatgt 5275  
|||||  
3574 ACTCAAAAGCGGTATATAGGTTATCCACGATATCAGGGGATTAACCCAGAAAGACATGT 3633  
|||||  
5276 gagcaaaagcgcaaaagcgcaagaaacgtaaaagcgcggttctgctgagctttcc 5335  
|||||  
3634 GAGCAAAAGCGCAAAAGCGCAAGAACCGTAAAAAGCGCGGTTGCTGGCGTTTTC 3693  
|||||  
5336 atagctccgccccctctacagacacacaaaatcgaagctcaagctcaagatgtgagca 5395  
|||||  
3694 ATAGGCTCCGCCCGCTGACGAGCATCAAAATCGAGCTCAAGTCAAGAGTGGCGAA 3753  
|||||  
5396 accgcagagacataaagataccagcgcttcccccgtgaagctccctcgtctgcctc 5455

|||||  
Db 3754 ACCGCAGAGACTATAAAGATACAGGCGTTCCCGCTGAGAGCTCCCTCGCGCTCC 3813  
|||||  
5456 ctgttcgaacctgcgcgtctccgatatccgtcgcgtctctcccttgaggagcgctg 5515  
|||||  
3814 GTTTCGACCCCTGCGGCTTACCGGATACGTGCTCGGCTTCTCCCTTGGGAAGCGCTGG 3873  
|||||  
5516 cgccttctaatgctccacgctgtaagtactcagttcgtgtaagtcgtctgctccaaac 5575  
|||||  
3874 CGCTTTCATAGCTCAGCGCTGATGATTCAGTGGGTGTAGTGGTGTGCTCCAGC 3933  
|||||  
5576 tgggctgtgtgcaagaccccccggtcagcccgagcgtcgtcgtctacgtaatactc 5635  
|||||  
3934 TGGGCTGTGTGACAGAACCCCGCTTACGCCGACCGCTGGCGCTTATCCGTTAACTATC 3993  
|||||  
5636 gctctgagtcacaacccggttaagacacagctatcgcacgtgagcgagcgacgttaac 5695  
|||||  
3994 GTTGTAGTCCAAACCCGGTAAACAGACTTATGCGCACTGGACAGCCACTGGTAACA 4053  
|||||  
5696 ggaatagcagagcgagatgtgagcggtgctacagagttcttgaagtgtgtgcttaact 5755  
|||||  
4054 GGATTAGCAGACGAGATGATGAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCTTAAC 4113  
|||||  
5756 acggtctacactagaagagcagatattgtatcctgcgtctcgtcgtcgtcgtcgtcgt 5815  
|||||  
4114 ACGGCTACACTAGAGAGACAGATTGTTGATCTGGCTCTCTCAAGCCACTTACCTTCG 4173  
|||||  
5816 gaaaaagagttgtgagctcttgatccggaacaaacacacgcgtgtgagcggtgtttt 5875  
|||||  
4174 GAAAAAGATTGTGAGCTTGTGATCTGGCAACAAACACCGCTGTGAGCGGTGTTT 4233  
|||||  
5876 ttgtttgaagcagcagatacgcgcgagaaaaaagaatcacaagaatccttgatct 5935  
|||||  
4234 TTGTTGCAAGCAGACAGATTACGGCAGAAAAAAGATCTCAAGAAAGATCTTGTATCT 4293  
|||||  
5936 ttctcaaggggtctgaagcgtcaagtgtgaaacgaacacgtcttaaggaatttggta 5995  
|||||  
4294 TTTTACGGGGGTGACCTCAGTGAAGCAAACTCACGTTAAGGATTTGGTCACTGA 4353  
|||||  
5996 gatitcaaaaagagatctcaacactagatccctttaaatlaaatagaattlaaactca 6055  
|||||  
4354 GATTTCAAAAAGATCTTCACTAGATCTTTTAAATTAATAAATGAAGTTTAAATCA 4413  
|||||  
6056 tctaaagatataatagtaacttgcgtcagttacaaatgcttaactcagtaggacac 6115  
|||||  
4414 TCTAAAGTATATATAGTAACCTTGCTGACAGTTACCAATGCTTATATCAGTGACAC 4473  
|||||  
6116 ctatctcagcgatcgtctatattggttcaatcaatagttgctgactcccggtgtgaga 6175  
|||||  
4474 CTATCTCAGCGATCTGTATTTGCTTATCTCATATGATGCTGACTCCCGCTCGTGA 4533  
|||||  
6176 taactacagataagagaggttaacatctggccccagtgctgcaatgataccggaac 6235  
|||||  
4534 TTACTACGATAGGAGGAGGCTTACCATCTGGCCCCAGTGTGCAATGATACCGGAAAC 4593  
|||||  
6236 cagctcaacggtctccagattatcaagcaataaacacgacccggaagggcgagcgca 6295  
|||||  
4594 CACGCTCACCGGCTCCAGATTATACGAATAAACACACCGCGGGAAGGCGCAGCGCA 4653  
|||||  
6296 gaagtgtctcggcaacttaccgcgtccatccaggtcattatattgttgcgggaacta 6355  
|||||  
4654 GAAGTGTCTCGCAACTTATATCCGCTTCATCCAGTGTATTAATTTGGCCGGAACCTA 4713  
|||||  
6356 gagtaagtagttcgcagttatagtttgcgaacggttgttgcattgtctacaagacatc 6415  
|||||  
4714 GAGTAAGTAGTTCCGCAATTATAGTTTGGCCAAAGTTTGGCATTTGCTACAGCATCG 4773  
|||||  
6416 tgggtcagcgtcgtcgttggtagtgcattacgaactccggttcccaacgatacaagc 6475  
|||||  
4774 TGGGTCAAGCTCGTGGTTGGTATGGCTTCATTACAGTCCGGTTCCCAAGATCAAGGC 4833  
|||||  
6476 gagttacatgatacccaatgttgtgcaaaaagcggttagctcctcgtctcccgatcg 6535  
|||||



Db 4834 GAGTTACATGATCCCCCATGTTGTCAGAAAAAGCGTTAGCTCCCTCCGTCCTCCGATCG 4893  
 QY 6536 ttgtcaagaatagtttgccgcgaagtggtatcactcaatggtatggaacgaatgcaat 6595  
 Db 4894 TTGTGAGAAATAGTATGCGCGCAGTGTATCATCTCATGTTATGCGCAGCTGCATATTT 4953  
 QY 6596 ccttactgcatgccaaccgaatgcttcttgtaagtggtgagtaaccgaat 6655  
 Db 4954 CTCTACTGTCATGCGCATCCGTAAGATGCTTTCTGTAAGTGGATGACTCAACCAAGT 5013  
 QY 6656 cacttgagaatagtgtatgctgcgcgaacgaatgctcttgccggcggtcaatacggata 6715  
 Db 5014 CATTTGTAGATAGTATGTCGGCGAGCGGTTGCTTCCCGCGTCATATCGGGATA 5073  
 QY 6716 ataccgcccacataagaagaacttaaaagtgtcatcatgtaaaacgttctcggggc 6775  
 Db 5074 ATACCGCGCCATAGCAGAACTTTAAAGTGTCTCATCATGGAAGAAAGCTTCTCGGGGC 5133  
 QY 6776 gaaactctcaagatcttcaacgctgttgagatccagtlcagatgaaccacgtgcac 6835  
 Db 5134 GAAAGCTCTCAAGATCTTACCGCTGTGAGATCCAGTTGATGATACCCACCTCGTGAC 5193  
 QY 6836 ccaactgacttcaagcatcttacttcaacgaagcttctggtgagcaaaaaacagaa 6895  
 Db 5194 CCAACTGATCTCAGCATCTTTACTTTCACGACGTTCTGGGTGAGCAAAAAACAGAA 5253  
 QY 6896 ggcgaatgcccgaataaagaagaagggcgacacggaatgtgaatactcatctc 6955  
 Db 5254 GGCAGAAATGCGCAAAAGGAAATAGGGCGACGGAATGTGAATCTATCTCT 5313  
 QY 6956 tcccttllcaatacttgaagcatlatalaaggttatgtlccatgagcgagatacat 7015  
 Db 5314 TCCTTTTCAATATATGAGACATTTATCAGGGTTATGTCTCATGAGCGGATACATAT 5373  
 QY 7016 ttgaatgattagaataaacaataaggggtccgcgcacattcccgaaaaagtc 7075  
 Db 5374 TTGAATGTATTTAGAAATTAACAAATAGGGGTCCGCGCACATTTCCCGGAAAAAGTGC 5433  
 QY 7076 caactgaagcttaagaacacattatatacgaacttaacataaaatagcgatata 7135  
 Db 5434 CACTTACGCTTAAGAAACCATTAATATCATGACATTAACCTATAAATAAGCGTATCA 5493  
 QY 7136 cgaagcccttgcgtcgcgcggttgcgtgatacaggtgaaaaactcgaacatgcagc 7195  
 Db 5494 CGAGGCCCTTTGCTCGCGCGCTTGGGTGATGACGTGAAGCCTGTGACACATGCAGC 5553  
 QY 7196 tcccggaagacgttcacagcttgcgtgaagcggaagccgggagcaagaacccgtcagg 7255  
 Db 5554 TCCCGGAGACGGTCAACAGCTTGTGTAAAGCGGATCCGGGAGCAGAACAGCCGTCAGG 5613  
 QY 7256 ggcgcgtcaagcggtgtggtgcgggtgtcgggggtggttaactatgagcaatacgaaga 7315  
 Db 5614 GCGCGTAGGGGGGTGGGGGTGGGGGTGGGCTTAACCTATGCGCATCAGACAGAA 5673  
 QY 7316 ttgtaatgagtgacacatalgctggtgtgaaataccgcaacagatggtgaagaagaat 7375  
 Db 5674 TTGTACTGAGAGTGCACCATATGCGGTGTGAATACCGCACATGCTTAAGAGAAAT 5733  
 QY 7376 accgcatcaggaatgtgtaacgttaatacttgttaaaatlcgcgttaaatltgtta 7435  
 Db 5734 ACCGCATCAGGAATGTAAACGTTATATTTGTTAAATTCGCGTTAAATTTTGTGA 5793  
 QY 7436 aatcagctcatltttaaccaatagcggaatcggaataatccctataaatcaaaaga 7495  
 Db 5794 AATCAGCTCATTTTAAACCAATAGCGCGAAATCGGCAAAATCCCTTATTAATCAAAAGA 5853  
 QY 7496 atagacgaagataggttgagtggttgcaggttgaacaagaagccactataaaga 7555  
 Db 5854 ATAGACCGAGATAGGTTAGTGTGTCTTCAGTTGGAACAAAGTCCACTATTAAGAA 5913  
 QY 7556 cgtgactccaacgltcaaaaggcgaaaaacgctctatcaaggcgatgcccactacgtga 7615  
 Db 5914 CGTGGACTCCAAAGCTCAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTAGTGA 5973

QY 7616 accatcaccttaactcaagtttltgggtgcgaagtgccgtfaaagcaactaatcgaaacc 7675  
 Db 5974 ACCATCACCTTAATCAAGTTTGGGGTGCAGTGCCGTAAGCATTAATGGAACCC 6033  
 QY 7676 taaaggagccccgaattagaacttgaacgagggaagccggaacgltgaggaaga 7735  
 Db 6034 TAAAGGAGCCCCGATTTAGACTTGACGGGGAAGCCGGCAACGTGGCGGAAGAAG 6093  
 QY 7736 agggaaagaagcgaaaggagcggtgaagcgctggaagtgcaagtgtgagcggtcagc 7795  
 Db 6094 AGGGAAGAAAGCGAAAGGAGCGGGCGCTAGGGCGCTGCGCAAGTGTAGCGGTACGCTGCG 6153  
 QY 7796 cgttaaccacacacccgcgcgttaatgagccgctacagaaggcggtgcgcacatcgcc 7855  
 Db 6154 CATTACCACACACCCCGCGCTTAATGCGCGCTACAGGGCGCTCGGCCATTGCGC 6213  
 QY 7856 attcaggctacgaactgttggaagggcgatcggtgcggcctcttcgtatlaagcca 7915  
 Db 6214 ATTCAAGGCTACGCACTGTTGGGAAGGGGATGCGTGGGGCTCTTGGCTATTAGCCA 6273  
 QY 7916 gctggc 7921  
 Db 6274 GGCTGC 6279

Search completed: November 18, 2001, 04:50:04  
 Job time: 35006 sec



**THIS PAGE BLANK (USPTO)**



12	4171.6	52.5	4275	9	AAAB80446	Modified factor VI
13	4170	52.5	4272	9	AAAB80447	Modified factor VI
14	4161	52.4	4373	20	AAAX82258	Beta-domain delete
15	4152.2	52.3	4830	9	AAAB81544	Human Factor VIII-
16	4108.8	51.7	4545	9	AAAB80444	Modified factor VI
17	3748	47.2	5035	18	AAAT69811	Factor VIII-48695
18	3708.6	46.7	4616	9	AAAB81545	Human Factor VIII-
19	3678	46.3	5094	21	AAA48231	DNA construct HSO,
20	3312	41.7	11846	20	AAAX82261	Factor VIII protein
21	3311	41.7	12022	20	AAAX82260	Factor VIII protein
22	3189.8	40.2	4334	19	AAAV12142	Homo sapiens factor
23	3189.8	40.2	4334	20	AAAX91195	Porcine factor VIII
24	3189.8	40.2	4334	22	AAAC96541	cDNA encoding porc
25	2999.4	37.8	6981	21	AAZ45931	Nucleotide sequenc
26	2985.2	37.6	6253	20	AAAX08454	AAV vector sequenc
27	2985.2	37.6	6280	20	AAAX08455	AAV vector sequenc
28	2985.2	37.6	6280	20	AAAX08456	AAV vector sequenc
29	2985.2	37.6	6565	21	AAZ45925	Nucleotide sequenc
30	2985.2	37.6	7054	21	AAZ45927	Nucleotide sequenc
31	2985.2	37.6	7405	21	AAZ45926	Nucleotide sequenc
32	2947	37.1	7492	21	AAZ45929	Nucleotide sequenc
33	2836.4	35.7	6142	21	AAZ45932	Nucleotide sequenc
34	2836.4	35.7	6142	21	AAZ45933	Nucleotide sequenc
35	2836.4	35.7	6142	21	AAZ45933	Nucleotide sequenc
36	2836.4	35.7	6714	21	AAZ45930	Nucleotide sequenc
37	2836.4	35.7	6924	21	AAZ45934	Nucleotide sequenc
38	2836.4	35.7	6924	21	AAZ45935	Nucleotide sequenc
39	2830.6	35.6	10930	16	AAAB81226	Plasmid pm16-1. C
40	2830.6	35.6	10950	16	AAAB81225	Plasmid pm16. Ch1
41	2794.2	35.2	4118	18	AAAT69188	Construct pGEM-hTR
42	2794.2	35.2	4118	18	AAAT69189	Construct pGEM-hTR
43	2794.2	35.2	7175	21	AAA26728	Plasmid pGAL1pN157
44	2794.2	35.2	7177	21	AAA25989	Plasmid pGAL1pN157
45	2762.6	34.8	6094	18	AAV05703	Plasmid pcre-LBD(g
ALIGNMENTS						
RESULT	1					
ID	AAAD00121					
	AAAD00121 standard; DNM; 11933 BP.					
XX	AAAD00121;					
DT	31-JUL-2000 (first entry)					
XX						
DE	Recombinant adeno associated vector construct, pAAV-F8-1.					
XX						
KW	Recombinant Adeno Associated Vector; rAAV; pAAV-F8-1; human Factor VIII;					
KW	hFVIII; hNF-3 albumin promoter; human elongation factor-1alpha; EF1alpha;					
KW	human growth hormone; hGH; inverted terminal repeat; ITR; haemophililla;					
KW	gene therapy; ds.					
XX						
OS	Adeno associated virus.					
XX						
PN	W0200023116-A1.					
XX						
PD	27-APR-2000.					
XX						
PF	19-OCT-1999; 99WC-US24495.					
XX						
PR	20-OCT-1998; 98US-0104994.					
PR	24-MAR-1999; 99US-0125974.					
PR	30-JUL-1999; 99US-0364862.					







Dh	1664	gcccccgatgacgaagtttataaagtcaatatattgaaacaatgccccccagcgattggt	1745
Qy	1737	aggaagatcaaaaaaagtcgcgattatctgcatcacagaatgaaaccttaagaactctgtgaa	1796
Dh	1744	aggaagtacaaaaaagtcgcgattatctgcatcacagaatgaaaccttaagaactctgtgaa	1803
Qy	1797	gctattcaagcatgaaatcaagaaatcttgggaacttactttagtgggaagtctggagaca	1856
Dh	1804	gctattcaagcatgaaatcaagaaatcttgggaacttactttagtgggaagtctggagaca	1863
Qy	1857	ctgtctgattatatttaagaaatcaagaacgaacataataacttaccctcaaggatc	1916
Dh	1864	ctgtctgattatatttaagaaatcaagaacgaacataataacttaccctcaaggatc	1923
Qy	1917	actgatagccgctccctctgtatacttaagagaatattacaanaagtgataaaaacatttgaagat	1976
Dh	1924	actgatagccgctccctctgtatacttaagagaatattacaanaagtgataaaaacatttgaagat	1983
Qy	1977	tttccaattctgcagagagaataattcaataataatgagacgtgacgtgtagaagatgag	2036
Dh	1984	tttccaattctgcagagagaataattcaataataatgagacgtgacgtgtagaagatgag	2043
Qy	2037	ccaacttaactcagatcctctggtgacgaccgctattactctgaattcttgtaaatgag	2096
Dh	2044	ccaacttaactcagatcctctggtgacgaccgctattactctgaattcttgtaaatgag	2103
Qy	2097	agaatacttaagcttcaagagacatctgtgcccctccccaactctgtctaaagaatctgtgat	2156
Dh	2104	agaatacttaagcttcaagagacatctgtgcccctccccaactctgtctaaagaatctgtgat	2163
Qy	2157	caaaagagaaaccagataaatgctcagaacaagagaatgctcactgtttctctgtatctgat	2216
Dh	2164	caaaagagaaaccagataaatgctcagaacaagagaatgctcactgtttctctgtatctgat	2223
Qy	2217	gagaaacggaagctcgtgatacctcccaagaagaaatatacaagctttctcccaatccagctcga	2276
Dh	2224	gagaaacggaagctcgtgatacctcccaagaagaaatatacaagctttctcccaatccagctcga	2283
Qy	2277	gtgacagcttgaagatccagaaggttcacaagctcccaacatcatgacaagacataatggtcat	2336
Dh	2284	gtgacagcttgaagatccagaaggttcacaagctcccaacatcatgacaagacataatggtcat	2343
Qy	2337	gtttttgatagttttgcaattgtgtcagtttgttcgataaggtgtagactggtacatctcta	2396
Dh	2344	gtttttgatagttttgcaattgtgtcagtttgttcgataaggtgtagactggtacatctcta	2403
Qy	2397	agcatctgagacacagagactgactctcttctgtcctctctctcgtatataactcttaaacac	2456
Dh	2404	agcatctgagacacagagactgactctcttctgtcctctctctcgtatataactcttaaacac	2463
Qy	2457	aaaaatgtctatgaagaacaacaacatcccatcccatctccagagagaacgtcttcagt	2516
Dh	2464	aaaaatgtctatgaagaacaacaacatcccatcccatctccagagagaacgtcttcagt	2523
Qy	2517	tcgaatgaaaaaacccaggtctatgaaatctctggtgtgcacaactcagaaatcttcgnaacaga	2576
Dh	2524	tcgaatgaaaaaacccaggtctatgaaatctctggtgtgcacaactcagaaatcttcgnaacaga	2583
Qy	2577	ggcatctgcgcctctactgaaaggttctctagtctgtgcacaagaacactggtgatatccag	2636
Dh	2584	ggcatctgcgcctctactgaaaggttctctagtctgtgcacaagaacactggtgatatccag	2643
Qy	2637	gacagttatgaagaatacttcagacataactctgtcgtgaaataaaaacatggccattgaaaccaaga	2696
Dh	2644	gacagttatgaagaatacttcagacataactctgtcgtgaaataaaaacatggccattgaaaccaaga	2703
Qy	2697	agcttctcccaagatctcaagacaccttgcagctatgagcaaaagcaatttaatgcacccca	2756
Dh	2704	agcttctcccaagatctcaagacaccttgcagctatgagcaaaagcaatttaatgcacccca	2709
Qy	2757	ccagttcttgaacgcgatcaacgcggaataactgcttactactcttcagtcagatccaag	2816
Dh	2710	ccagttcttgaacgcgatcaacgcggaataactgcttactactcttcagtcagatccaag	2745

QY	2817	gaatttgactatgatgatataccatatacagtttgaaatgaaagaaagatttggacattat	2876
Db	2746	gaattgactatgatgatataccatatacagtttgaaatgaaagaaagatttggacattat	2805
QY	2877	gattgaagatgaaaatcagaagcccccgaagcttcaaaaagaaaacagcactatttatt	2936
Db	2806	gattgagatgaaaatcagaagcccccgaagcttcaaaaagaaaacagcactatttatt	2865
QY	2937	gctgcagatggaagagctcctgagatataatgagatgagtaagctccccaagtctlaagaac	2996
Db	2866	gctgcagatggaagagctcctgagatataatgagatgagtaagctccccaagtctlaagaac	2925
QY	2997	aagggtcagatgagcagtggtccctcaagtccaagaagtgtgtttccagaattactgat	3056
Db	2926	aagggtcagatgagcagtggtccctcaagtccaagaagtgtgtttccagaattactgat	2985
QY	3057	ggctccttttaccagaacctataacgctggagaactcaaatgaaacattttggagctctggg	3116
Db	2986	ggctccttttaccagaacctataacgctggagaactcaaatgaaacattttggagctctggg	3045
QY	3117	ccatataaagcagcgaagtltgaaatatataatcattggttccatccaagccct	3176
Db	3046	ccatataaagcagcgaagtltgaaatatataatcattggttccatccaagccct	3105
QY	3177	gctcccatctccttattctagaaccttattcttatgaggaatcagaagcagaagca	3236
Db	3106	gctcccatctccttattctagaaccttattcttatgaggaatcagaagcagaagca	3165
QY	3237	gaacctagaanaaactttgtccaagcttaatgaaaccaaacttacttttggaaagtga	3296
Db	3166	gaacctagaanaaactttgtccaagcttaatgaaaccaaacttacttttggaaagtga	3225
QY	3297	catcataatgacccacataaagatgagtgtgaactgacgcaagcctggtcttattcctgat	3356
Db	3226	catcataatgacccacataaagatgagtgtgaactgacgcaagcctggtcttattcctgat	3285
QY	3357	gtttgacctggaanaaagatgtgtgaccccaagccctgatttgaaacctctgtctgcacact	3416
Db	3286	gtttgacctggaanaaagatgtgtgaccccaagccctgatttgaaacctctgtctgcacact	3345
QY	3417	aacacactgaaccccgctcatatgaggaacaaagtgcatacagatagaagaaattgtgcctgttttc	3476
Db	3346	aacacactgaaccccgctcatatgaggaacaaagtgcatacagatagaagaaattgtgcctgttttc	3405
QY	3477	accatcctttgatgagaccaaagaactgtgtacttactctgtaaaatataatgaaagaactgcag	3536
Db	3406	accatcctttgatgagaccaaagaactgtgtacttactctgtaaaatataatgaaagaactgcag	3465
QY	3537	gtctcctcgaatatatacagatggaataatcccaacttttaagaagaattatcggtttcatgca	3596
Db	3466	gtctcctcgaatatatacagatggaataatcccaacttttaagaagaattatcggtttcatgca	3525
QY	3597	atccaaatgctataataatgatacaactaccctggtcttagtaaatggtccagaatccaagaatt	3656
Db	3526	atccaaatgctataataatgatacaactaccctggtcttagtaaatggtccagaatccaagaatt	3585
QY	3657	cgaatggtatctgtccagcatatggcagcgaatgaaacatccatctcatattcaagtga	3716
Db	3586	cgaatggtatctgtccagcatatggcagcgaatgaaacatccatctcatattcaagtga	3645
QY	3717	catgtgttcaactgtacgaaaaaaaagaagatataaagtgcacgtgtcaactcttatcca	3776
Db	3646	catgtgttcaactgtacgaaaaaaaagaagatataaagtgcacgtgtcaactcttatcca	3705
QY	3777	gattgttttttgagacagttggaatgttaccacatccaagaactgtgaatttggcgggttgaaatgc	3836
Db	3706	gattgttttttgagacagttggaatgttaccacatccaagaactgtgaatttggcgggttgaaatgc	3765
QY	3837	cttatttgcagacatctacatagctgtggatgagcacactttttctgtgtgtacagcaataag	3896
Db	3766	cttatttgcagacatctacatagctgtggatgagcacactttttctgtgtgtacagcaataag	3825



[illegible][illegible]



Db 5753 cttaagtataatagtaaaactgtgtcagacgttaaccatgctaaatcagtgagccacc 5812  
QY 6117 tatctagcgatcgtctatcttctgtcatccatagcttgcttgatcccgctcgtgtgat 6176  
Db 5813 tatccagcgatccgtctatcttctgtcatccatagcttgcttgatcccgctcgtgtgat 5872  
QY 6177 aactacgatacggagggccttaaccatctggcccagctgcgtgcgaatgataccgcagagccc 6236  
Db 5873 aactacgatacggagggccttaaccatctggcccagctgcgtgcgaatgataccgcagagccc 5932  
QY 6237 acgctcacccggtctccagattatcatcagaacaacccagccgagggccgagccgag 6296  
Db 5933 acgctcacccggtctccagattatcatcagaacaacccagccgagggccgagccgag 5992  
QY 6297 aagaggtcccgcaactttatccgctccatccagctctatcaattgtgtccgaggagctag 6356  
Db 5993 aagaggtcccgcaactttatccgctccatccagctctatcaattgtgtccgaggagctag 6052  
QY 6357 agtaagtagtccgcaacttaatagtttctgcgaacggttctgtccatctgtcacaagcactcgt 6416  
Db 6053 agtaagtagtccgcaacttaatagtttctgcgaacggttctgtccatctgtcacaagcactcgt 6112  
QY 6417 ggtgtcagcgtcgtctgtgtgtatgcttcacacagctccggtcccaacgataaagcg 6476  
Db 6113 ggtgtcagcgtcgtctgtgtgtatgcttcacacagctccggtcccaacgataaagcg 6172  
QY 6477 agtatacgtatcccccattgtgtgcgaataaagcggttagctccctcgtctccatccgt 6536  
Db 6173 agtatacgtatcccccattgtgtgcgaataaagcggttagctccctcgtctccatccgt 6232  
QY 6537 tgtcagaagaagaattgtgcgcgcaagtgtatatacctcatgtatgtgcagcaactgtcataatc 6596  
Db 6233 tgtcagaagaagaattgtgcgcgcaagtgtatatacctcatgtatgtgcagcaactgtcataatc 6292  
QY 6597 tcttactgctatgcacccgtaagaatgcttctgtgcacggtgagtaactccaacaaagtc 6656  
Db 6293 tcttactgctatgcacccgtaagaatgcttctgtgcacggtgagtaactccaacaaagtc 6352  
QY 6657 attctggaatagtgtatgtgcgcgcaacgagattgctctgtgcgcggtccatatacggagtaa 6716  
Db 6353 attctggaatagtgtatgtgcgcgcaacgagattgctctgtgcgcggtccatatacggagtaa 6412  
QY 6717 tacccgcgcacatagcagaactttaaaggtctcatcatctgtgaaacgcttcttcggcg 6776  
Db 6413 tacccgcgcacatagcagaactttaaaggtctcatcatctgtgaaacgcttcttcggcg 6472  
QY 6777 aaaaactccaagatcttaaccgctgtgtgagatccagttcgatgtaaccactcgtgcacc 6836  
Db 6473 aaaaactccaagatcttaaccgctgtgtgagatccagttcgatgtaaccactcgtgcacc 6532  
QY 6837 caactgactctcagcactcttactctcaccaagcttctcgggtgagcaaaaacaggaag 6896  
Db 6533 caactgactctcagcactcttactctcaccaagcttctcgggtgagcaaaaacaggaag 6592  
QY 6897 gcaaaaatgcgcgaaaaagggaaataaggcgcaacgcgaaaatgtctgaaatactactctt 6956  
Db 6593 gcaaaaatgcgcgaaaaagggaaataaggcgcaacgcgaaaatgtctgaaatactactactctt 6652  
QY 6957 ccttttcaatatatgtgaagcattatcaggggttatgtctcatatgagcggaatacatt 7016  
Db 6653 ccttttcaatatatgtgaagcattatcaggggttatgtctcatatgagcggaatacatt 6712  
QY 7017 tgaatgattatgaaaaataaacaataaggggttcgcgcgacattcccccgaaaaagggcc 7076  
Db 6713 tgaatgattatgaaaaataaacaataaggggttcgcgcgacattcccccgaaaaagggcc 6772  
QY 7077 acctgacgtctaaagaaacattatcatcagatataaccataaataagcgatcac 7136  
Db 6773 acctgacgtctaaagaaacattatcatcagatataaccataaataagcgatcac 6832  
QY 7137 gaaggccttctcgtctcgcgcgttctggtgtatgacggtgaaacactctgacacatgcact 7196  
Db 6833 gaaggccttctcgtctcgcgcgttctggtgtatgacggtgaaacactctgacacatgcact 6892

QY 7197 cccggagacggtcacacagcttctgtgtaagcggatgcccggagacgaacagcccgtaagg 7256  
Db 6893 cccggagacggtcacacagcttctgtgtaagcggatgcccggagacgaacagcccgtaagg 6952  
QY 7257 cgcgtcagcgggtgtgtgcgggtgtcggggcgtgtccttaactatgcggatcagaagcagat 7316  
Db 6953 cgcgtcagcgggtgtgtgcgggtgtcggggcgtgtccttaactatgcggatcagaagcagat 7012  
QY 7317 tgtactgaagagtcaacatatgtcgtgtgaaataccgcacagatgtcgtlaaggaanaata 7376  
Db 7013 tgtactgaagagtcaacatatgtcgtgtgaaataccgcacagatgtcgtlaaggaanaata 7032  
QY 7377 ccgcatacgaagaattgtgaacggttaatatgttgttaaaatcgcgttaaatgtttgttaa 7436  
Db 7033 -----aaatgtgaacggttaatatgttgttaaaatcgcgttaaatgtttgttaa 7082  
QY 7437 atcagctcatcttcttaaccataagccgaatcgcgaataatccctataataataaagaa 7496  
Db 7083 atcagctcatcttcttaaccataagccgaatcgcgaataatccctataataataaagaa 7142  
QY 7497 taagccgagataagggttgagtgctgtccagtttggaacaagatccactataaagaac 7556  
Db 7143 taagccgagataagggttgagtgctgtccagtttggaacaagatccactataaagaac 7202  
QY 7557 gtggactccaagctcaaaaggcgaaaacgcttatcagaagcgatggccccactagtgaa 7616  
Db 7203 gtggactccaagctcaaaaggcgaaaacgcttatcagaagcgatggccccactagtgaa 7262  
QY 7617 ccatcaccccttaacaaagtcttctgggggtcgaagggtgccgtlaaagcactaaatcggaaacct 7676  
Db 7263 ccatcaccccttaacaaagtcttctgggggtcgaagggtgccgtlaaagcactaaatcggaaacct 7222  
QY 7677 aaaggagagcccccgtattgaagcttgaacggggaagccggcggaacgtgtgcgagaaagaa 7736  
Db 7323 aaaggagagcccccgtattgaagcttgaacggggaagccggcggaacgtgtgcgagaaagaa 7382  
QY 7737 ggggaagaagcgaagaagagcgcgctagaggcgcttggaaggtgtaacggtcacagctgcgc 7796  
Db 7383 ggggaagaagcgaagaagagcgcgctagaggcgcttggaaggtgtaacggtcacagctgcgc 7442  
QY 7797 gtaaccacacaccccgcgcttaatgtcgcgcgtacacagggcgct 7842  
Db 7443 gtaaccacacaccccgcgcttaatgtcgcgcgtacacagggcgct 7488

RESULT 2  
AAx82259  
ID AAx82259 standard; cDNA; 9164 BP.  
AC  
XX  
AC AAx82259;  
XX  
DT 18-AUG-1999 (first entry)  
XX  
DE Beta-domain deleted Factor VIII protein encoding gene (uncorrected).  
XX  
KW Factor VIII protein; gene modification; gene therapy; clinical disorder;  
KW Splicing pattern; RNA processing; gene regulation; beta-domain; human;  
KW ss.  
XX  
OS Homo sapiens.  
XX  
PN W09929848-A1.  
PD 17-JUN-1999.  
XX  
XX 25-NOV-1998; 98WO-US25354.  
PF  
XX 16-JAN-1998; 98US-0071596.  
PR 05-DEC-1997; 97US-0067614.  
XX  
XX (IMMU-) IMMUNE RESPONSE CORP.  
XX







QY 2095 agagagatctagcttcagagactcatctggccctccctcatctgtctacaagaatctgag 2154  
|||||  
Db 2681 agagagatctagcttcagagactcatctggccctccctcatctgtctacaagaatctgag 2740  
QY 2155 atcaaaaggaacacagataatgtcgaagaaggaatgtcatccctgtttctgtatttg 2214  
|||||  
Db 2741 atcaaaaggaacacagataatgtcgaagaaggaatgtcatccctgtttctgtatttg 2800  
QY 2215 atgggaacggagctgtgacctcaccagagaatatacaagctttctccccaatccagctg 2274  
|||||  
Db 2801 atggaacggagctgtgacctcaccagagaatatacaagctttctccccaatccagctg 2860  
QY 2275 gaggcagctgagagatccagaagttccaagctccaacatcatgacaagcatcaatgct 2334  
|||||  
Db 2861 gaggcagctgagagatccagaagttccaagctccaacatcatgacaagcatcaatgct 2920  
QY 2335 atgttttgatagtttgcaagttgtcagtttggttgcatgaggtgcatctgtacattc 2394  
|||||  
Db 2921 atgttttgatagtttgcaagttgtcagtttggttgcatgaggtgcatctgtacattc 2980  
QY 2395 taagcatgtggagcacaagctgacttctctgtctctctctctcgtgatatccctcaaac 2454  
|||||  
Db 2981 taagcatgtggagcacaagctgacttctctgtctctctctcgtgatatccctcaaac 3040  
QY 2455 acaaaatgctctatgaaagacacactcaaccctattcccatcttcagggagaactgtcttca 2514  
|||||  
Db 3041 acaaaatgctctatgaaagacacactcaaccctattcccatcttcagggagaactgtcttca 3100  
QY 2515 tgcgatgtgaaaaacccaagtcatagtatctctggtgtgccaactcagacttcggaaca 2574  
|||||  
Db 3101 tgcgatgtgaaaaacccaagtcatagtatctctggtgtgccaactcagacttcggaaca 3160  
QY 2575 gaggcagcagccgcttactgtgaagtttctgtgtgtgacagaagaacacgtgtgattatcag 2634  
|||||  
Db 3161 gaggcagcagccgcttactgtgaagtttctgtgtgtgacagaagaacacgtgtgattatcag 3220  
QY 2635 aggcagctatgaaagacatcttcacagatctgtcgtgaataaacaatgccaattgaacaa 2694  
|||||  
Db 3221 aggcagctatgaaagacatcttcacagatctgtcgtgaataaacaatgccaattgaacaa 3280  
QY 2695 gaagctctcccaagaattcaagacacccctagcactaggaacaaagttaatgccacc 2754  
|||||  
Db 3281 gaagctctcccaaga-----cc 3298  
QY 2755 caccagcttgaaaacgcgcatacaagcgaataaactgttactactcttcagtcagaatcaag 2814  
|||||  
Db 3299 caccagcttgaaaacgcgcatacaagcgaataaactgttactactcttcagtcagaatcaag 3358  
QY 2815 aggaatctgactatgatatccatcaagttgaatggaagaagaatttgacatt 2874  
|||||  
Db 3359 aggaatctgactatgatatccatcaagttgaatggaagaagaattttgacatt 3418  
QY 2875 atgataaggaatgaaaatcagagccccgcagctttcaaaaagaaaacgcagacatattta 2934  
|||||  
Db 3419 atgataaggaatgaaaatcagagccccgcagctttcaaaaagaaaacgcagacatattta 3478  
QY 2935 ttgctgtcaggtggaagaagctctcgtggtatatagtgagtagctccccaatgtttcaagaa 2994  
|||||  
Db 3479 ttgctgtcaggtggaagaagctctcgtggtatatagtgagtagctccccaatgtttcaagaa 3558  
QY 2995 acaaggtcagagatggtcaggtgtccctcagttccaagaagttgtttccaagaattttatctg 3054  
|||||  
Db 3539 acaaggtcagagatggtcaggtgtccctcagttccaagaagttgtttccaagaattttatctg 3598  
QY 3055 atggctccttactcagcccttataccgttggaagaactaaatgaaacatttggagctcttg 3114  
|||||  
Db 3599 atggctccttactcagcccttataccgttggaagaactaaatgaaacatttggagctcttg 3658  
QY 3115 ggcacatatagaagcagaagtttgaagataataatcatatcgtgtaactttcagaatcagagct 3174  
|||||  
Db 3659 ggcacatatagaagcagaagtttgaagataataatcatatcgtgtaactttcagaatcagagct 3718

QY 3175 ctgcctccatcttctctatcttagcttacttcttaatgaaagatcagagcagaag 3234  
|||||  
Db 3719 ctgcctccatcttctctatcttagcttacttcttaatgaaagatcagagcagaag 3778  
QY 3235 cagaacctagaaaaaacttgttcaagcctcaatgaaacaaaacttacttvtggaagtgc 3294  
|||||  
Db 3779 cagaacctagaaaaaacttgttcaagcctcaatgaaacaaaacttacttvtggaagtgc 3838  
QY 3295 aacatcatatgcccacactaaagaatgattttagctgtcaagcctgtggtatttctctg 3354  
|||||  
Db 3839 aacatcatatgcccacactaaagaatgattttagctgtcaagcctgtggtatttctctg 3898  
QY 3355 atgttgaacctgaaaaaagatgtgacacacagcctgatttggacccctctgtgtccaca 3414  
|||||  
Db 3899 atgttgaacctgaaaaaagatgtgacacacagcctgatttggacccctctgtgtccaca 3958  
QY 3415 ctaacacactgaaacctgtctcatggtggaacaagtgtacaagtacaggaattgtctgtttt 3474  
|||||  
Db 3959 ctaacacactgaaacctgtctcatggtggaacaagtgtacaagtacaggaattgtctgtttt 4018  
QY 3475 tcacacatctttagatgagacccaaaagctggttacttctcactgtgaataatagtgaagaactgca 3534  
|||||  
Db 4019 tcacacatctttagatgagacccaaaagctggttacttctcactgtgaataatagtgaagaactgca 4078  
QY 3535 ggcctccctgtcaatccagaatggaagatcccaacttttaagaagaattatcgcttccatg 3594  
|||||  
Db 4079 ggcctccctgtcaatccagaatggaagatcccaacttttaagaagaattatcgcttccatg 4138  
QY 3595 caatcaatggtctacataatgatatcactaccctgtgttaagtgtgtcagatcaagaaga 3654  
|||||  
Db 4139 caatcaatggtctacataatgatatcactaccctgtgttaagtgtgtcagatcaagaaga 4198  
QY 3655 ttcgatgttatctgtcagacatggcgacgaatggaataatccatcttcatattcatcttg 3714  
|||||  
Db 4199 ttcgatgttatctgtcagacatggcgacgaatggaataatccatcttcatattcatcttg 4258  
QY 3715 gacatgtcttcaactgtacgaaaaaaagaggtataaatgtgcaactgtacatctctac 3774  
|||||  
Db 4259 gacatgtcttcaactgtacgaaaaaaagaggtataaatgtgcaactgtacatctctac 4318  
QY 3775 caagtgctttttgagaacgttgaataatgttacaatccaaagcttgaatttgcgggtggaat 3834  
|||||  
Db 4319 caagtgctttttgagaacgttgaataatgttacaatccaaagcttgaatttgcgggtggaat 4378  
QY 3835 gacctatttggcgaacatctcatctgtgtggaatgacaacttttctgtgtgacagcaata 3894  
|||||  
Db 4379 gacctatttggcgaacatctcatctgtgtggaatgacaacttttctgtgtgacagcaata 4438  
QY 3895 agtctcagactcccttgggaatgtctctgtgacacatctagaatttccagatcaagct 3954  
|||||  
Db 4439 agtctcagactcccttgggaatgtctctgtgacacatctagaatttccagatcaagct 4498  
QY 3955 caaggacatatggaacagtgggcccccaagctgtgccaagactcattatctccgatalca 4014  
|||||  
Db 4499 caaggacatatggaacagtgggcccccaagctgtgccaagactcattatctccgatalca 4558  
QY 4015 atgcctgtgaacacaaagagcccttttctgtgatacaaggtgtgactgttggcaccaatga 4074  
|||||  
Db 4559 atgcctgtgaacacaaagagcccttttctgtgatacaaggtgtgactgttggcaccaatga 4618  
QY 4075 ttattcagcgaatcaagaacccaggtgtcccgtaagaagttctccagcctcaatctctc 4134  
|||||  
Db 4619 ttattcagcgaatcaagaacccaggtgtcccgtaagaagttctccagcctcaatctctc 4678  
QY 4135 agttatcatcatgtatagcttctgtatggtgaagaagtgtgcaagacttalcgaaggaattcca 4194  
|||||  
Db 4679 agttatcatcatgtatagcttctgtatggtgaagaagtgtgcaagacttalcgaaggaattcca 4738  
QY 4195 ctggaaccttaatggtcttcttggcgaatgtgatctcatctgggataaacaacatatctt 4254  
|||||  
Db 4739 ctggaaccttaatggtcttcttggcgaatgtgatctcatctgggataaacaacatatctt 4798  
QY 4255 ttaacctccaattattgtctcgatataatccgttttgcaacccaactcattatagatttgca 4314



Db	4799	ttaaacctccaattatctgtcgtacacatccgttttcacccaactaatatagatctgcga	4858
QY	4315	gcaactcttcgcacatgagatgtgaatggcgttgatattaaatagttgcagacatgccaattgggaa	4374
Db	4859	gcaactcttcgcacatgagatgtgaatggcgttgatattaaatagttgcagacatgccaattgggaa	4918
QY	4375	tgagaaataagcaataatcaaatgacagatattactgtctcaactcttaacttaccataatgct	4434
Db	4919	tgagaaataagcaataatcaaatgacagatattactgtctcaactcttaacttaccataatgct	4978
QY	4435	ttgcacctgtgtccctctcaaaaagctcgcacttccactcccaaggaggaataatgtcctgga	4494
Db	4979	ttgcacctgtgtccctctcaaaaagctcgcacttccactcccaaggaggaataatgtcctgga	5038
QY	4495	gaactcaggtggaataatccaaaagagtgctgcgaagtggacttccagaagaacatgtaag	4554
Db	5039	gaactcaggtggaataatccaaaagagtgctgcgaagtggacttccagaagaacatgtaag	5098
QY	4555	tcacagaagtaactactacagaggaataaatctctgtaccacagcatgtatgtgaaagagct	4614
Db	5099	tcacagaagtaactactacagaggaataaatctctgtaccacagcatgtatgtgaaagagct	5158
QY	4615	ttccatctccacagcagtlcaagaatgagccatcagtgagctctctttttccaagaatggcnaag	4674
Db	5159	ttccatctccacagcagtlcaagaatgagccatcagtgagctctctttttccaagaatggcnaag	5218
QY	4675	ttaaagtttttccaagggaataataagaatctctccacacccgttgtagactctcagaaccac	4734
Db	5219	ttaaagtttttccaagggaataataagaatctctccacacccgttgtagactctcagaaccac	5278
QY	4735	cgttaccgactcgtctacactcttgaaattcacccccacagagttgggtgcacacagatctgcctga	4794
Db	5279	cgttaccgactcgtctacactcttgaaattcacccccacagagttgggtgcacacagatctgcctga	5338
QY	4795	ggatggagaggtcttggtgctgcagagcacagagacacttactgactcgaagcagatctcttcga	4854
Db	5339	ggatggagaggtcttggtgctgcagagcacagagacacttactgactgaggtgcgcacatgcagacc	5398
QY	4855	gggga-----	4859
Db	5399	tgccacgcgcgtcaactctctccctcctcactcagctccacagggcaggtccctccctgcttgct	5458
QY	4860	-----	4859
Db	5459	tctacacttgctgaataatccacagacagacatgcgtcttgaaagctctcgtatataatcat	5518
QY	4860	-----	4859
Db	5519	cagtcctgcacttctcttggtgctggggccagagaggtgcataccaatttaacttaactctcta	5578
QY	4860	-----	4871
Db	5579	ccattttctgcagctgctccagattactctcttccctccaataataactagagcaaaaaga	5638
QY	4872	agtcagaataaa-----	4883
Db	5639	agtcagaataaa-----	5698
QY	4884	-----	4883
Db	5699	caactcctcgtgtgtagaaaacctgtagtaagaaccttgaaaaaagatatatgtatgt	5758
QY	4884	-----	4883
Db	5759	aactgttctatctgagcttataatggttatacaataaagaacatagcatcacaaatttcaca	5818
QY	4884	-----	4883
Db	5819	aataaagcatttttccaactgcattctagttgtgtgtccaaaactacaatgatact	5878
QY	4884	-----	4883



QY	5418	ccagcgcttcccccctggaagctccctcgtctgcctctccgtctccgaacctgcgcgttac	54177
Db	7019	ccaaagcgttcccccttggaaagctccccctcgtctgcctctccgtctccgaacctgcgcgttac	7078
QY	5478	cggaatactctccgcctcttcctccctccgggaagcgtgcggccttccaaatctaagctg	5537
Db	7079	cggaatacgtctccgcctcttccctccctccgggaagcgtgcggccttccaaatctaagcgtg	7138
QY	5538	taagatcatcagatctcgtgtgaagtcgtctccccaagctgcggcgtctgtgcacgaacccc	5597
Db	7139	taagatcatcagatctcgtgtgaagtcgtctccccaagctgcggcgtctgtgcacgaaccccc	7198
QY	5598	cggttcaagcccggaacgcgtgcgcctctacccgtaactatcgtctcttggtcccaacccgtgaag	5657
Db	7199	cggttcaagcccggaacgcgtgcgcctctacccgtaactatcgtctcttggtcccaacccgtgaag	7258
QY	5658	acaacgaattatcgcacactgtgcagccagccacactgtgaacaagatctgaagcgaagatgct	5717
Db	7259	acaacgaattatcgcacactgtgcagccagccacactgtgaacaagatctgaagcgaagatgct	7318
QY	5718	agcgcggtgtctacagagttcttgaagtgtgtgcgcctaacctacgcgtctacactagaagaagct	5777
Db	7319	agcgcggtgtctacagagttcttgaagtgtgtgcgcctaacctacgcgtctacactagaagaagct	7378
QY	5778	attggtatctgcgcctcgtctgtgaagccaaattacacctgcgaaagaagctgttgtaaccttg	5837
Db	7379	attggtatctgcgcctcgtctgtgaagccaaattacacctgcgaaagaagctgttgtaaccttg	7438
QY	5838	atcccgcaacaacaacacacgcgtgtgtagcgtgtgtttttgttttgtaacagcagcagatcac	5897
Db	7439	atcccgcaacaacaacacacacgcgtgtgtagcgtgtgtttttgttttgtaacagcagcagatcac	7498
QY	5898	gcgcgagaagaaaaaggtctctcaagaagatactccttgatctcttctcaagggtctgcagctca	5957
Db	7499	gcgcgagaagaaaaaggtctctcaagaagatactccttgatctcttctcaagggtctgcagctca	7558
QY	5958	gtgcgaacgaanaactccacgcttlaaggggtttttgtgtctatgagatctacaagaagatctcac	6017
Db	7559	gtgcgaacgaanaactccacgcttlaaggggtttttgtgtctatgagatctacaagaagatctcac	7618
QY	6018	ctagaacctctttaaattaaataatgaagttttaataatcaatctaagtatatatgaagtaaac	6077
Db	7619	ctagaacctctttaaattaaataatgaagttttaataatcaatctaagtatatatgaagtaaac	7678
QY	6078	ttgtctcgaagttacccaatagttacttaagtgtagggcacctctcaagagctctctact	6137
Db	7679	ttgtctcgaagttacccaatagttacttaagtgtagggcacctctcaagagctctctact	7738
QY	6138	tcgttatccatagttgtgcgtactccccgtctcgtgtgtagataaactacgatacggagggcgt	6197
Db	7739	tcgttatccatagttgtgcgtactccccgtctcgtgtgtagataaactacgatacggagggcgt	7798
QY	6198	accatctgcggcccaagtgctgcgaatgtatgtagccgggaagcccaagctcacacgcgtccaagatt	6257
Db	7799	accatctgcggcccaagtgctgcgaatgtatgtagccgggaagcccaagctcacacgcgtccaagatt	7858
QY	6258	atcagcaataaacaagccagccagcgaagggcccgagcgcgaagctgtgtcttcgaacttacc	6317
Db	7859	atcagcaataaacaagccagccagcgaagggcccgagcgcgaagctgtgtcttcgaacttacc	7918
QY	6318	cgccctcatccagttcatataattgtgtgcgggaagcttagagtgtagttctgcgaagttaa	6377
Db	7919	cgccctcatccagttcatataattgtgtgcgggaagcttagagtgtagttctgcgaagttaa	7978
QY	6378	taagtttcggaagcgttgtgtgcacgtctacaggaactcgtgtgtgtcaagctcgtctgttg	6437
Db	7979	taagtttcggaagcgttgtgtgcacgtctacaggaactcgtgtgtgtgtcaagctcgtctgttg	8038
QY	6438	tatggtcttaactcagctccggttctcccaagatcaagcgcgagttatagatcccccagt	6497
Db	8039	tatggtcttaactcagctccggttctcccaagatcaagcgcgagttatagatcccccagt	8098

QY	6458	gtgcacaaaagcgggtttagctctccgtccgtccgcgtgtgtcagaagaatgtgtgcgcg	6537
Db	8099	gtgcacaaaagcgggtttagctctccgtccgtccgcgtgtgtcagaagaatgtgtgcgcg	8156
QY	6558	agtgttcatcactaaatgttttagggagacacgtcaatcttccttaccgttcacatccgt	6612
Db	8159	agtgttcatcactaaatgttttagggagacacgtcaatcttccttaccgttcacatccgt	8218
QY	6618	aagatgcttctctgttagctgtgtgatactaaccaagtcattctgtgaaatgttatgcg	6677
Db	8219	aagatgcttctctgttagctgtgtgatactaaccaagtcattctgtgaaatgttatgcg	8276
QY	6678	gcgaccgagatgtgctctgtcccgcgctcaatcagggataaacccgcgcacatagacaac	6737
Db	8279	gcgaccgagatgtgctctgtcccgcgctcaatcagggataaacccgcgcacatagacaac	8338
QY	6738	tttaaaagtgtctcaatctttagaaaacgttcttggggcgaaaacttcaagatcttacc	6797
Db	8339	tttaaaagtgtctcaatctttagaaaacgttcttggggcgaaaacttcaagatcttacc	8396
QY	6798	gctgtctgagatccagcttcgatgttaaccactctgtgacccaactgtatcttcagcatctt	6857
Db	8399	gctgtctgagatccagcttcgatgttaaccactctgtgacccaactgtatcttcagcatctt	8455
QY	6858	tactttcaacagcgcttctctgggtgtgagcaaaaacaggaagycaaaatgtccgcaaaaagg	6917
Db	8459	tactttcaacagcgcttctctgggtgtgagcaaaaacaggaagycaaaatgtccgcaaaaagg	8518
QY	6918	aataagggcgcagcaaggaaatgttgaaatccatcacctcttccttccaatatattggaag	6977
Db	8519	aataagggcgcagcaaggaaatgttgaaatccatcacctcttccttccaatatattggaag	8578
QY	6978	cattatccaaggcttatgtgtccatgagcggatatacatatttgaattgatttagaaaaataa	7037
Db	8579	cattatccaaggcttatgtgtccatgagcggatatacatatttgaattgatttagaaaaataa	8633
QY	7038	acaatatagggtgttcccgcgacattctcccgcaaaaagtcgacacgtgactaagaacaacat	7097
Db	8639	acaatatagggtgttcccgcgacattctcccgcaaaaagtcgacacgtgactaagaacaacat	8696
QY	7098	tattatcatgtacattaaacctataaaatagcgctatcaagagcccttcgtctcgcgcg	7157
Db	8699	tattatcatgtacattaaacctataaaatagcgctatcaagagcccttcgtctcgcgcg	8756
QY	7158	tttcggtgatacgggttaaaaacctctaacacatgtgcagcgtcccggaagacgttcaagcttg	7217
Db	8759	tttcggtgatacgggttaaaaacctctaacacatgtgcagcgtcccggaagacgttcaagcttg	8818
QY	7218	tctgttaagcggatgtcccgggagcagacaagcccgctcaaggcgctgtctcagcggtgttgcg	7277
Db	8819	tctgttaagcggatgtcccgggagcagacaagcccgctcaaggcgctgtctcagcggtgttgcg	8878
QY	7278	gtgtccggggtcgtgcttctaactatgtcggtcatagaagcagattgttactgtagaggtgcaccatat	7337
Db	8879	gtgtccggggtcgtgcttctaactatgtcggtcatagaagcagattgttactgtagaggtgcaccatat	8938
QY	7338	gcgggtgtgaaaaatccagcagatgctgtaaaggaaaaataccgcatcagg	7386
Db	8939	gcgggtgtgaaaaatccagcagatgctgtaaaggaaaaataccgcatcagg	8987
RESULT 3			
AAD00122			
ID	AAD00122 standard; DNA: 4999 BP.		
XX	AAD00122:		
XX			
XX	31-Jul-2000 (first entry)		
DE	Recombinant adeno associated vector construct, pvm4.1c6f8deltab.		
XX			
Recombinant Adeno Associated Vector: IAAV; pvm4.1c6f8deltab; promoter; human Factor VIII; hFVIII; EF1alpha; human elongation factor-1alpha;			







QY	1393	ttctactgttttctgtcatactcttccccaacaaatgtagcatggaagcttatgtcaaaag	1452
Db	1389	ttctactgttttctgtcatactcttccccaacaaatgtagcatggaagcttatgtcaaaag	1448
QY	1453	tagacaagctgtccagaggaagacccaactacgaatgaaaaataatgaaagcggaaagact	1512
Db	1449	tagacaagctgtccagaggaagacccaactacgaatgaaaaataatgaaagcggaaagact	1508
QY	1513	atgaaatgaaatcttactgatacttctgaaatggaatgtagtgcaggtttgatatgacaacttc	1572
Db	1509	atgaaatgaaatcttactgatacttctgaaatggaatgtagtgcaggtttgatatgacaacttc	1568
QY	1573	cttcctttatccaatattgcgtcgaagtgtgccaagaagcaltcccaaaacttgggttacatbca	1632
Db	1569	cttcctttatccaatattgcgtcgaagtgtgccaagaagcaltcccaaaacttgggttacatbca	1628
QY	1633	tttgtctgtgaaagaggaagactctggagctatgctcccttaagttccctgcgccccgaatgacaga	1692
Db	1629	tttgtctgtgaaagaggaagactctggagctatgctcccttaagttccctgcgccccgaatgacaga	1688
QY	1693	gttataaaagcatatatttgaaacaatggtcccttcagcgagatgtgttagagaaatgacaaaaag	1752
Db	1689	gttataaaagcatatatttgaaacaatggtcccttcagcgagatgtgttagagaaatgacaaaaag	1748
QY	1753	tcagattatgcatbacaagaatgaaacctttaagactcgtgaaagctatctcaagcatgaaat	1812
Db	1749	tcagattatgcatbacaagaatgaaacctttaagactcgtgaaagctatctcaagcatgaaat	1808
QY	1813	cagaatcttgggaaccttacttatagggggaagtgtgagacacacactgttgaattatatta	1872
Db	1809	cagaatcttgggaaccttacttatagggggaagtgtgagacacacactgttgaattatatta	1868
QY	1873	agaatcaagcaagcaagccataatacatctacccctcaagagatacactgatagtccgtcctt	1932
Db	1869	agaatcaagcaagcaagccataatacatctacccctcaagagatacactgatagtccgtcctt	1928
QY	1933	tgatccaagagaattaccaaaagggtgaaacaatttgaaagatcttccaattccgcacag	1992
Db	1929	tgatccaagagaattaccaaaagggtgaaacaatttgaaagatcttccaattccgcacag	1988
QY	1993	gagaataattcaaatataaatctgacaagtgtactgtagaagaatctgggccaactaaatcaagatc	2052
Db	1989	gagaataattcaaatataaatctgacaagtgtactgtagaagaatctgggccaactaaatcaagatc	2048
QY	2053	ctcgggtgcctgaccccggtattactactagtttccgttaatatgtagagagaattactaactcaag	2112
Db	2049	ctcgggtgcctgaccccggtattactactagtttccgttaatatgtagagagaattactaactcaag	2108
QY	2113	gactcatatggccctctcctcactcgtctacaagaagatctgtatgatacacaagaggaaccaga	2172
Db	2109	gactcatatggccctctcctcactcgtctacaagaagatctgtatgatacacaagaggaaccaga	2168
QY	2173	taatgtcagaacaagaagaatgtcatccctgtttctgttataattgtatgtagaaccgaaagctggt	2232
Db	2169	taatgtcagaacaagaagaatgtcatccctgtttctgttataattgtatgtagaaccgaaagctggt	2228
QY	2233	acctccaagaagaataatacaagcttcttccccaatccaagctcgagggtctgaagctttagagatc	2292
Db	2229	acctccaagaagaataatacaagcttcttccccaatccaagctcgagggtctgaagctttagagatc	2288
QY	2293	cagaagctccaagcctccaacaactcgcacagaacaaatgtagtctgtattttagatgttgc	2352
Db	2289	cagaagctccaagcctccaacaactcgcacagaacaaatgtagtctgtattttagatgttgc	2348
QY	2353	agttgtcaagtttctgtcatgaggttgtagatactgtgtacatcttaagcatctggagcacaga	2412
Db	2349	agttgtcaagtttctgtcatgaggttgtagatactgtgtacatcttaagcatctggagcacaga	2408
QY	2413	ctgactcccttctgtcttcttctctgtagatataccttccaacacaaaatggtctatgaag	2472
Db	2409	ctgactcccttctgtcttcttctctgtagatataccttccaacacaaaatggtctatgaag	2468
QY	2473	acaaactcaaccatactccatcttccagagagaactgtctcatgtctgatgtaaaaccag	2532

---

Db	2469	aaacaactcaaccatactccatcttccagaggaagaaactgtcttcatgtcaggtgtaaaaccag	2528
QY	2533	gtctatggaatctctgggtgtgcacaactcagactctcggaaacagagcatgacgccttac	2592
Db	2529	gtctatggaatctctgggtgtgcacaactcagactctcggaaacagagagcatgacgccttac	2588
QY	2593	tgaaggttcttagttgtgacaagaacactggtgtgattatagagagacaggttatgaaagta	2652
Db	2589	tgaaggttcttagttgtgacaagaacactggtgtgattatagagagacaggttatgaaagta	2648
QY	2653	tttcagcatctctgtctgtaaaaaacaatgcatgtgaaccacaagaagcttccccaagaatt	2712
Db	2649	tttcagcatctctgtctgtaaaaaacaatgcatgtgaaccacaagaagcttccccaagaatt	2707
QY	2713	caagaacaccttagacactagcgcaaaagcaatttaatgccaaccaccagttcttgaagcgc	2772
Db	2708	-----ccaccagttcttgaagcgc	2726
QY	2773	atcaagcggaaataactcgttactactcttcagtcagatcaagagaaattgactatgag	2832
Db	2777	atcaagcggaaataactcgttactactcttcagtcagatcaagagaaattgactatgag	2786
QY	2833	ataccatactcagttgaaatgaaagagagaattttgacattatgatatgagatgaaatc	2892
Db	2787	ataccatactcagttgaaatgaaagagagaattttgacattatgatatgagatgaaatc	2846
QY	2893	agaagcccgagagcttccaanaagaaacaacagacataatttatgtctgcagttggagagc	2952
Db	2847	agaagcccgagagcttccaanaagaaacaacagacataatttatgtctgcagttggagagc	2906
QY	2953	ttctgggattatgggattgtagtgcctcccaatggttcttaagaacagggctcaagatggca	3012
Db	2907	ttctgggattatgggattgtagtgcctcccaatggttcttaagaacagggctcaagatggca	2966
QY	3013	gtgtcccccagttccaagaagaattgttttccaggaatttaacgtatggtcccttaactcagc	3072
Db	2967	gtgtcccccagttccaagaagaattgttttccaggaatttaacgtatggtcccttaactcagc	3026
QY	3073	ccctataccgttggagaaactaaatgaaacatttgggaatccctcggggccatataaagacag	3132
Db	3027	ccctataccgttggagaaactaaatgaaacatttgggaatccctcggggccatataaagacag	3086
QY	3133	aagttgaaagataatactatgtagtaactttcagaagaatcagagcctctgtcccatcttctct	3192
Db	3087	aagttgaaagataatactatgtagtaactttcagaagaatcagagcctctgtcccatcttctct	3146
QY	3193	attcctagccttattcttataatgagaagatcagaggtcgaagagcaagaactgaaaaaaact	3252
Db	3147	attcctagccttattcttataatgagaagatcagaggtcgaagagcaagaactgaaaaaaact	3206
QY	3253	ttgtcaagcctaattgaaaccaaaacttacttttggaaagtgcacaatcatatgtgaccca	3312
Db	3207	ttgtcaagcctaattgaaaccaaaacttacttttggaaagtgcacaatcatatgtgaccca	3266
QY	3313	ctaaagatgaatttgactgcgcaaaagccttgggttatcttctcgtatggttgaacttgaaag	3372
Db	3267	ctaaagatgaatttgactgcgcaaaagccttgggttatcttctcgtatggttgaacttgaaag	3326
QY	3373	atgtgcactcagagcctgatttggaaacccctctggtctgcacaaactaaacacatgaaacctg	3432
Db	3327	atgtgcactcagagcctgatttggaaacccctctggtctgcacaaactaaacacatgaaacctg	3386
QY	3433	ctcatvggagacaagtgcaggtacagaagaattgctcgttttttccaacatctttagatgaga	3492
Db	3387	ctcatvggagacaagtgcaggtacagaagaattgctcgttttttccaacatctttagatgaga	3446
QY	3493	ccaaagctggttacttgaacttgaataatggaanaaaacttgaaaggtctcccgcatatcc	3552
Db	3447	ccaaagctggttacttgaacttgaataatggaanaaaacttgaaaggtctcccgcatatcc	3506
QY	3553	agatggaagatcccaactttaaagagaattatcgccttccaatgacaaatgtgcacataa	3612



Db	3507	agatggaagatcccaacttttaagaagattaacgtcccaagcaatacaatgctcaactaa	3508
QY	3613	tggatacactactgtgcttagtaattggctccagataccaagatctgaatgtatctgtcca	3612
Db	3567	tggatacactactgtgcttagtaattggctccagataccaagatctgaatgtatctgtcca	3626
QY	3673	gcattggcagcaatgaanaacatccattctattcaatttcagtgagcaatggttcaacttac	3732
Db	3627	gcattggcagcaatgaanaacatccattctattcaatttcagtgagcaatggttcaacttac	3686
QY	3733	gaaaaaaggaggagataaanaatggcactgtatactctatccacagtggtttttaaagcaag	3792
Db	3687	gaaaaaaggaggagataaanaatggcactgtatactctatccacagtggtttttaaagcaag	3746
QY	3793	tggaaatgttaccataccaagcttggaattcttgccgggttggaatgtgcattatlgccagcatc	3852
Db	3747	tggaaatgttaccataccaagcttggaattcttgccgggttggaatgtgcattatlgccagcatc	3806
QY	3853	tacatgtctgggaatgaagcaactttttctggtgtacagcaataagtgtaacatccctctg	3912
Db	3807	tacatgtctgggaatgaagcaactttttctggtgtacagcaataagtgtaacatccctctg	3866
QY	3913	gaattgctcttcggcacattagaagaatttccagattcaagcttcaaggagcaatatagagact	3972
Db	3867	gaattgctcttcggcacattagaagaatttccagattcaagcttcaaggagcaatatagagact	3926
QY	3973	gggcccacaagcttgccagagctcatattatccggatccaatacaatgaccttgagccacaag	4032
Db	3927	gggcccacaagcttgccagagctcatattatccggatccaatacaatgaccttgagccacaag	3986
QY	4033	agcccttttcttgatacaaggttgatctctgtgccaacaatgatattatcaagcatcaaga	4092
Db	3987	agcccttttcttgatacaaggttgatctctgtgccaacaatgatattatcaagcatcaaga	4046
QY	4093	cccaagggtgcccgtcagagaagtctcccaagccctcacatctccagtttatactcaatgata	4152
Db	4047	cccaagggtgcccgtcagagaagtctcccaagccctcacatctccagtttatactcaatgata	4106
QY	4153	gtcttgaatgggaagaagtggcagagcttatacgaagaaatccacatcgagaccttaattggtct	4212
Db	4107	gtcttgaatgggaagaagtggcagagcttatacgaagaaatccacatcgagaccttaattggtct	4166
QY	4213	tctttggcaatgfggattcatctctggataaaacacaataatttttaaccctccaattatctg	4272
Db	4167	tctttggcaatgfggattcatctctggataaaacacaataatttttaaccctccaattatctg	4226
QY	4273	ctcgatatcatccggtttgcaaccaactatataatgatttggagagacctcttcgaatgagat	4332
Db	4227	ctcgatatcatccggtttgcaaccaactatataatgatttggagagacctcttcgaatgagat	4286
QY	4333	tgaatggcttggaatttaataatgttgcagcatgycatctgggaatggagatgaagaacaat	4392
Db	4287	tgaatggcttggaatttaataatgttgcagcatgycatctgggaatggagatgaagaacaat	4346
QY	4393	cagatgcacagattactgcttcatcactacttlaacaataatgtttgccaacctggtctcctt	4452
Db	4347	cagatgcacagattactgcttcatcactacttlaacaataatgtttgccaacctggtctcctt	4406
QY	4453	caaaagcttgcaacttcaactccaaaggggagtaattgctctggagacccctcagtgtaataatc	4512
Db	4407	caaaagcttgcaacttcaactccaaaggggagtaattgctctggagacccctcagtgtaataatc	4466
QY	4513	caaaaagcttgcttgcaagtggacttccagaagaacaatggaagtacaggaatcaactctc	4572
Db	4467	caaaaagcttgcttgcaagtggacttccagaagaacaatggaagtacaggaatcaactctc	4526
QY	4573	agggaagtaaaaactctgcttaccagcaatgtaatgtaaggaggttccatctccagcaatc	4632
Db	4527	agggaagtaaaaactctgcttaccagcaatgtaatgtaaggaggttccatctccagcaatc	4586
QY	4633	aagatgagcaatcagtggaactctctttttccaagaatggccaagaagtatttcaaggaa	4692
Db	4587	aagatgagcaatcagtggaactctctttttccaagaatggccaagaagtatttcaaggaa	4646

OY	4693	atuaagactccttaacaccgttggtgtagaacctctaagaaccacccggttacctgaactcgctaac	4752
OY	4693	atuaagactccttaacaccgttggtgtagaacctctaagaaccacccggttacctgaactcgctaac	4752
Db	4647	atcaagactccttaacaccgttggtgtagaacctctaagaaccacccggttacctgaactcgctaac	4706
OY	4753	ttcgaattcacccccagaagtgggtgcaccagattgccctgaagatgagggtttctggct	4812
Db	4707	ttcgaattcacccccagaagtgggtgcaccagattgccctgaagatgagggtttctggct	4766
OY	4813	gcgagcgacaggaactctactgactcgagcgagttcttctgaggggatacgacaataaaa	4872
Db	4767	gcgagcgacaggaactctactgactcgagc-----ctataaag	4805
OY	4873	gacggaataaaacgcacagcgggtgtgtgggttgtttgttcgatccagatctaggaaaccccta	4932
Db	4806	gaaatattttctcatgcaatagtggtgtgtttttgtgtgtggcgcgaggaaccccta	4865
OY	4933	gtgatggagtggcgaactcccctctgcgcgcctgcctcactaagtcgacgcgccgggca	4992
Db	4866	gtgatggagtggcgaactcccctctgcgcgcctgcctcactaagtcgacgcgccgggca	4925
OY	4993	aagcccgcgctcgcggaactctgtgtcgcgcgcgcctcaagtgaagcgagcgcgag	5051
Db	4926	aagtcgcgcgcagcgccggcgttcttcgcgcgcgcgcctcaagtgaagcgagcgcgag	4984
 RESULT 4 AAx88293 ID AAx88293 standard; DNA; 4629 BP. XX AC AAx88293; XX DT 24-SEP-1999 (first entry) XX DE Human Factor VIII with B domain deleted cDNA. XX KW Adenoviral vector; Factor VIII; Factor IX; clotting factor; treatment; KM haemostatic; haemophilia A; haemophilia B; gene therapy; ss. XX OS Homo sapiens. XX PN USS935935-A. XX PD 10-AUG-1999. XX PE 07-JUN-1995; 95US-0484891. XX PR 07-JUN-1995; 95US-0484891. PR 10-JUN-1993; 93US-0074920. PR 25-MAR-1994; 94US-0218335. XX PA (GENE-) GENETIC THERAPY INC. XX PI Connelly S, Kaleko M, Smith T; XX DR WPI; 1999-457617/38. XX PT Adenoviral vectors useful for treating hemophilia XX PS Example 1; Column 53-58; 90pp; English. XX CC This invention describes novel adenoviral vectors comprising at least CC one DNA sequence encoding a clotting factor (Factor IX or Factor VIII). CC The vectors of the invention have haemostatic activity. The vectors are CC useful for the treatment of hemophilia A or hemophilia B by gene therapy. CC This sequence represents human Factor VIII cDNA which has the B domain CC deleted. XX SQ Sequence 4629 BP; 1319 A; 1050 C; 1015 G; 1245 T; 0 other;			
 Query Match 55.6%; Score 4413; DB 20; Length 4629; Best Local Similarity 99.8%; Pred No. 0;			



Matches 4419; Conservative 0; Mismatches 10; Indels 0; Gaps 0;				
QY	420	atgcaaatagagctctcaccctgtcttctctgcttgcctttgagattctcttaagcc	479	
Db	1	atgcaaatagagctctcaccctgtcttctctgcttgcctttgagattctcttaagcc	60	
QY	480	accagaagatactacccctgggtgacagtgcacatgcagggacataatgcacaatgatctc	539	
Db	61	accagaagatactacccctgggtgacagtgcacatgcagggacataatgcacaatgatctc	120	
QY	540	ggtgagctgctctgagcagcaagaattctctctagagtgccaaatctttccatccaac	599	
Db	121	ggtgagctgctctgagcagcaagaattctctctagagtgccaaatctttccatccaac	180	
QY	600	accctgactctgtacaaaagactctgttgtgaattccaggttccacctttccaatc	659	
Db	181	accctgactctgtacaaaagactctgttgtgaattccaggttccacctttccaatc	240	
QY	660	gctaaagcgaagccacccttgatgggtctgtaggtccatccagctgagctgttat	719	
Db	241	gctaaagcgaagccacccttgatgggtctgtaggtccatccagctgagctgttat	300	
QY	720	gatacagtggtcattacatacttaagaacatggtctccatccatctgtcagttccatgctgt	779	
Db	301	gatacagtggtcattacatacttaagaacatggtctccatccatctgtcagttccatgctgt	360	
QY	780	ggtgtacctactggaagagctctcgaaggagctgaaatgatatgataccagatccaaag	839	
Db	361	ggtgtacctactggaagagctctcgaaggagctgaaatgatatgataccagatccaaag	420	
QY	840	gaagaagaagatgataaagctctccctggtggaagccatcatatgctgtgagctctg	899	
Db	421	gaagaagaagatgataaagctctccctggtggaagccatcatatgctgtgagctctg	480	
QY	900	aaagagaatgtccaatggtccttgaccacatgtgccttaacctatacttctccat	939	
Db	481	aaagagaatgtccaatggtccttgaccacatgtgccttaacctatacttctccat	540	
QY	960	gtggaactgtgtaaaagactgaaattcagcctatgtgaacctactagatgtagaagaa	1019	
Db	541	gtggaactgtgtaaaagactgaaattcagcctatgtgaacctactagatgtagaagaa	600	
QY	1020	gggaagctggtccaaagaaagaacagaccttgcaacaatttatactactttgtcgtta	1079	
Db	601	gggaagctggtccaaagaaagaacagaccttgcaacaatttatactactttgtcgtta	660	
QY	1080	tttgaatgaagggaaaagtgtgacatccagaacaaagaacctcttgatgcaggaatggat	1139	
Db	661	tttgaatgaagggaaaagtgtgacatccagaacaaagaacctcttgatgcaggaatggat	720	
QY	1140	gtcgcatactgtcgtggccttggtctaaatgcacacagttcaatgttataatgaacaggtct	1199	
Db	721	gtcgcatactgtcgtggccttggtctaaatgcacacagttcaatgttataatgaacaggtct	780	
QY	1200	ctgcacagttctgattggaatgcacaggaataatcagttcatgtgcatgtgatatgggc	1259	
Db	781	ctgcacagttctgattggaatgcacaggaataatcagttcatgtgcatgtgatatgggc	840	
QY	1260	accactctcgaagtgcataataatctcctcgaaggtccacacatttctctgtggaagaacct	1319	
Db	841	accactctcgaagtgcataataatctcctcgaaggtccacacatttctctgtggaagaacct	900	
QY	1320	cgcacagcgctcttggaataatctgcacaataacttctccttaactctcaaaaacctgtatg	1379	
Db	901	cgcacagcgctcttggaataatctgcacaataacttctccttaactctcaaaaacctgtatg	960	
QY	1380	gaacttggaacagttctactgttttgcataatctctccacaacaatgataggcatgtaa	1439	
Db	961	gaacttggaacagttctactgttttgcataatctctccacaacaatgataggcatgtaa	1020	
QY	1440	gcttatgtcaaaatagaacagctgtccaaggaagccccaactaagaatgaaaaatagaa	1499	
Db	1021	gcttatgtcaaaatagaacagctgtccaaggaagccccaactaagaatgaaaaatagaa	1080	

QY	1500	gaagcggaaagactatgatatgattcttactgattctgaatgtgatatggtcgaagtttgat	1559	
Db	1081	gaagcggaaagactatgatatgattcttactgattctgaatgtgatatggtcgaagtttgat	1140	
QY	1560	gatgacaactctctctcccttatccaatctcgatccaatctgcagaagacatcccaaaact	1619	
Db	1141	gatgacaactctctctcccttatccaatctcgatccaatctgcagaagacatcccaaaact	1200	
QY	1620	tgggtacattacatctgctctgtaaggaaggaactgggaactatgctcccttaagtcctgcc	1679	
Db	1201	tgggtacattacatctgctctgtaaggaaggaactgggaactatgctcccttaagtcctgcc	1260	
QY	1680	cccgatgacagaaggtataaaatccaattttgaacaatgagccctcagcgattgtagg	1739	
Db	1261	cccgatgacagaaggtataaaatccaattttgaacaatgagccctcagcgattgtagg	1320	
QY	1740	aagtacaaaaaagttccgattatggtacacacagatgaaacctttaagactcgtgaagct	1799	
Db	1321	aagtacaaaaaagttccgattatggtacacacagatgaaacctttaagactcgtgaagct	1380	
QY	1800	atccagcatgaaatccaggaatcttgggaaccttaacttatggggaagtgtgagacacactg	1859	
Db	1381	atccagcatgaaatccaggaatcttgggaaccttaacttatggggaagtgtgagacacactg	1440	
QY	1860	ttgattatattgaagatcaagcaagcagacacataacatctacccctccaggaatcact	1919	
Db	1441	ttgattatattgaagatcaagcaagcagacacataacatctacccctccaggaatcact	1500	
QY	1920	gattccctctcttctgtaattcaaggaatlaaccaaaggtgtlaaacaatttgaaggaattt	1979	
Db	1501	gattccctctcttctgtaattcaaggaatlaaccaaaggtgtlaaacaatttgaaggaattt	1560	
QY	1980	ccaattctgcaggaagaaatataatcaaatggaactgactgtagaagatgggcca	2039	
Db	1561	ccaattctgcaggaagaaatataatcaaatggaactgactgtagaagatgggcca	1620	
QY	2040	actaaatcagatctctggtccgtgacccggtatattactctgattctgttaatatggaagaa	2099	
Db	1621	actaaatcagatctctggtccgtgacccggtatattactctgattctgttaatatggaagaa	1680	
QY	2100	gattcagttccaggaactattggtccctctctcatctctgcacaaagaatctgtgatcaaa	2159	
Db	1681	gattcagttccaggaactattggtccctctctcatctctgcacaaagaatctgtgatcaaa	1740	
QY	2160	agaggaacacagataatgtlcaagaacaggaatgtcatcctgtttctgtattgtatgag	2219	
Db	1741	agaggaacacagataatgtlcaagaacaggaatgtcatcctgtttctgtattgtatgag	1800	
QY	2220	aaacgaagctgtgtacatccacagaagaatatacaacgctttctcccaattcagctgtgagtg	2279	
Db	1801	aaacgaagctgtgtacatccacagaagaatatacaacgctttctcccaattcagctgtgagtg	1860	
QY	2280	caagcttgagatccagaagtccaaagctcccaacatcatgtacagcatlcaaigtgctatgt	2339	
Db	1861	caagcttgagatccagaagtccaaagctcccaacatcatgtacagcatlcaaigtgctatgt	1920	
QY	2340	tttgaatagtttgagttgtcagttgtttgcatgaggtgtgcatcaactgtgatacttcaagc	2399	
Db	1921	tttgaatagtttgagttgtcagttgtttgcatgaggtgtgcatcaactgtgatacttcaagc	1980	
QY	2400	attggagcacaagatgaactccttctgctcttctctccgtgatatccctcaaacaaaa	2459	
Db	1981	attggagcacaagatgaactccttctgctcttctctccgtgatatccctcaaacaaaa	2040	
QY	2460	atgtgtcatgaaagacaactacacccatctccatctcaggaagaaactgtctcatgtctg	2519	
Db	2041	atgtgtcatgaaagacaactacacccatctccatctcaggaagaaactgtctcatgtctg	2100	
QY	2520	atggaaaaccacaggtctatagattctgggggtgccaacatccaaactttggaaacagagc	2579	
Db	2101	atggaaaaccacaggtctatagattctgggggtgccaacatccaaactttggaaacagagc	2160	



QY	2580	atgaacgccttactggaaggttcttaagtcttgagacaagaacactggttgatatctaacgagac	2639
Db	2161	atgacgcgccttactggaaggttctctagcttgagacaagaacactggttgatatctaacgagac	2220
QY	2640	agttatgaagaatctttcagcaactcttgctgagtaaaaaaatgtgcatttgaaaccaagaagc	2639
Db	2221	agttatgaagaatctttcagcaactcttgctgagtaaaaaaatgtgcatttgaaaccaagaagc	2280
QY	2700	ttctccagaatattcaagacacccctagcaacttgagcaaaagcaatttaatgcccaccacca	2759
Db	2281	ttctccagaatattcaagacacccctagcaacttgagcaaaagcaatttaatgcccaccacca	2340
QY	2760	gtcttgaaacgccaatcaacgggaaataactgcatactcttcagtcagatcaagaagaa	2819
Db	2341	gtcttgaaacgccaatcaacgggaaataactgcatactcttcagtcagatcaagaagaa	2400
QY	2820	attgcaactatgatatcacatatcagtttgaaatggaagaaggaaatttgacaatttatgat	2879
Db	2401	attgcaactatgatatcacatatcagtttgaaatggaagaaggaaatttgacaatttatgat	2460
QY	2880	gaagatgaanaatcaagagcccccgcagcttctcaaaagaanaacagacatatatttatgct	2939
Db	2461	gaagatgaanaatcaagagcccccgcagcttctcaaaagaanaacagacatatatttatgct	2520
QY	2940	gcaagtggagagcgtctgagatctaatgggactgaatgcctccaccaatggttctaagaacag	2999
Db	2521	gcaagtggagagcgtctgagatctgaatgggactgaatgcctccaccaatggttctaagaacag	2580
QY	3000	gtccagagatggcaggtgctccctcagttcaagaagttgtttccaggaatttactgagtc	3059
Db	2581	gtccagagatggcaggtgctccctcagttcaagaagttgtttccaggaatttactgagtc	2640
QY	3060	tccttctaacagcccttatcacgctggagagacttaataaactttggagactcctggggcca	3119
Db	2641	tccttctaacagcccttatcacgctggagagacttaataaactttggagactcctggggcca	2700
QY	3120	tatataagagcagaagtttgaagaataatacatggttaacttcaganaatcaagcctctcgt	3179
Db	2701	tatataagagcagaagtttgaagaataatacatggttaacttcaganaatcaagcctctcgt	2760
QY	3180	ccctattccttctatcttaaccttatatttcttatgagaagaatcaggaagcgaagcgagaa	3259
Db	2761	ccctattccttctatcttaaccttatatttcttatgagaagaatcaggaagcgagcgagaa	2820
QY	3240	ccctagaanaaaactttgccaagcctaatagaaaccaaaacttaactttggaagctgacaacat	3299
Db	2821	ccctagaanaaaactttgccaagcctaatagaaaccaaaacttaactttggaagctgacaacat	2880
QY	3300	catatgagcaccacaataaagaatgaatttgactgcaaaagcctgggcttaattctctgatggt	3359
Db	2881	catatgagcaccacaataaagaatgaatttgactgcaaaagcctgggcttaattctctgatggt	2940
QY	3360	gaacctggaaaaaagatgagcaactcagggccttgataggaccctctctggtctgcacataaac	3419
Db	2941	gaacctggaaaaaagatgagcaactcagggccttgataggaccctctctggtctgcacataaac	3000
QY	3420	aaccttgaacccctgctcatctggagacaagctagaagtaacaggaaattgctcgttttcaacc	3479
Db	3001	aaccttgaacccctgctcatctggagacaagctagaagtaacaggaaattgctcgttttcaacc	3060
QY	3480	atctcttgatgagacccaanaagctggtaacttctactgaactgaanaataatgnaaagaaactgcagggct	3539
Db	3061	atctcttgatgagacccaanaagctggtaacttctactgaactgaanaataatgnaaagaaactgcagggct	3120
QY	3540	cccttgacaatatccagatggaagaatcccaattttaagaagaatatgcgtctccatgcaatc	3599
Db	3121	cccttgacaatatccagatggaagaatcccaattttaagaagaatatgcgtctccatgcaatc	3180
QY	3600	aaatggctatactaaatggatcaacttaacccgggttgatgaatgctcaagggatcaagaattcga	3659
Db	3181	aaatggctatactaaatggatcaacttaacccgggttgatgaatgctcaagggatcaagaattcga	3240
QY	3660	tggatctgctcagcatggtggcagcaatgnaaaacatccatctatctaatctcaatggagcat	3719

Db	3241	tggtac	ctgcctcagcagatgaggagcaatggaataacatccattctattcttcagtggacat	3300
Qy	3720	ggttcac	ctgtacgaaaaaagagagatataaaatgagcactgtacacatctccatccagct	3779
Db	3301	gtgtccact	gtacgaaaaaagagagatataaaatgagcactgtacacatctccatccatccagct	3360
Qy	3780	gtttcttgaca	agatggaatgttaccatccaaagctgtgaattgtgaggtgtggaatgtcct	3839
Db	3361	gtttcttgaca	agatggaatgttaccatccaaagctgtgaattgtgaggtgtggaatgtcct	3420
Qy	3840	atttgcgag	catctacatgtctgggataagcacacttttctgtgtacagcaataatgtct	3899
Db	3421	atttgcgag	catctacatgtctgggataagcacacttttctgtgtacagcaataatgtct	3480
Qy	3900	cagactccccc	gggaatgtgctctcgtgagcacatttagagatttccagatttccagcttcagga	3959
Db	3481	cagactccccc	gggaatgtgctctcgtgagcacatttagagatttccagatttccagcttcagga	3540
Qy	3960	caata	tggacagatgggcccccaagcttgcacagacttcatattccgataccaatgtcc	4019
Db	3541	caata	tggacagatgggcccccaagcttgcacagacttcatattccgataccaatgtcc	3600
Qy	4020	tggagacc	caaagagcccttctctgtgataaggttgatctgtgtgcaccaatgtattt	4079
Db	3601	tggagacc	caaagagcccttctctgtgataaggttgatctgtgtgcaccaatgtattt	3660
Qy	4080	cacggc	catccaagaccagaggtgtgcctcctcagaagtcttccagctctaacctccatgct	4139
Db	3661	cacggc	catccaagaccagaggtgtgcctcctcagaagtcttccagctctaacctccatgct	3720
Qy	4140	atcaca	ctgtatagtctgtgatggaagaagtgtgcagacttaccgaggaattccactgtga	4199
Db	3721	atcaca	ctgtatagtctgtgatggaagaagtgtgcagacttaccgaggaattccactgtga	3780
Qy	4200	accttaat	tgctctctcttgcaatgtggaatcactggaataaacaacaataattttaac	4259
Db	3781	accttaat	tgctctctctcttgcaatgtggaatcactggaataaacaacaataattttaac	3840
Qy	4260	cctccaat	tattgtgtgatataacatccgttttcacccaacatattagactttgcagact	4319
Db	3841	cctccaat	tattgtgtgatataacatccgttttcacccaacatattagactttgcagact	3900
Qy	4320	cttcgca	ctgagctgtgactgtgactgtgatttaataagcttgacagatgccattggaagtgaag	4379
Db	3901	cttcgca	ctgagctgtgactgtgactgtgatttaataagcttgacagatgccattggaagtgaag	3960
Qy	4380	agtaaa	gcaataatcagatgcagattactgtcttcaacttccacttacaatatgtttgc	4439
Db	3961	agtaaa	gcaataatcagatgcagattactgtcttcaacttccacttacaatatgtttgc	4020
Qy	4440	acctgtgt	ctctcttcaaaagctcgacatctcacttccacttccaaaggagatgaatgtccgtgagact	4499
Db	4021	acctgtgt	ctctcttcaaaagctcgacatctcacttccacttccaaaggagatgaatgtccgtgagact	4080
Qy	4500	caggtga	atlaatccaaaagctgtgctgcaagtgaattccaaagacaaatgaaatgcaca	4559
Db	4081	caggtga	atlaatccaaaagctgtgctgcaagtgaattccaaagacaaatgaaatgcaca	4140
Qy	4560	ggagta	actactcagggagtaaaatctctgtcttaccagcagtatgtggaaggttctctc	4619
Db	4141	ggagta	actactcagggagtaaaatctctgtcttaccagcagtatgtggaaggttctctc	4200
Qy	4620	attcttcag	cagatcaagtgtgcatactcgtgactcttttttcagaatggaagaatgaag	4679
Db	4201	attcttcag	cagatcaagtgtgcatactcgtgactcttttttcagaatggaagaatgaag	4260
Qy	4680	gttttcag	gggaataatcaagactccttcaaacctgtgtgaactctctagaccacagctga	4739
Db	4261	gttttcag	gggaataatcaagactccttcaaacctgtgtgaactctctagaccacagctga	4320
Qy	4740	ctgactcgt	actccttgatctaccaccccaagtttgggtgcacagattgtcccttgagagt	4799



Db	4380	ctgactgcctactcttgaaatcacccccccagagttggtgacacaaattgacctgagatg	4380
Qy	4800	gaaggtctggcctcgagcgacacagacactctactgactcgagcgagttc	4848
Db	4381	gaagttctggcctcgagcgacacagacactctactgaggtggtgcactgc	4429
RESULT 5			
AA773164	standard; cDNA; 9354 BP.		
AA773164;			
08-APR-1998	(first entry)		
XX	cDNA encoding human B-domain deleted factor VIII.		
XX			
XX			
XX			
KW	Post-translational regulatory element; PRE: enhancer II; intronless gene;		
KW	surface antigen gene; cytoplasmic accumulation; targeted delivery;		
KW	near consensus splice sequence; blood coagulation factor; factor VIII;		
KW	factor IX; ss.		
XX			
OS	Homo sapiens.		
XX			
Key	Location/Qualifiers		
FT	CDS	2965..7380	
FT		/*tag= a	
FT		5165..5174	
FT	misc_feature	/*tag= b	
FT		/note= "5", near consensus site"	
FT		5695..5703	
FT	misc_feature	/*tag= c	
FT		/note= "5", near consensus site"	
FT		6320..6328	
FT	misc_feature	/*tag= d	
FT		/note= "5", near consensus site"	
FT		6595..5603	
FT	misc_feature	/*tag= e	
FT		/note= "5", near consensus site"	
FT		7045..7053	
FT	misc_feature	/*tag= f	
FT		/note= "5", near consensus site"	
FT		7143..7152	
FT	misc_feature	/*tag= g	
FT		/note= "5", near consensus site"	
FT		3296..3312	
FT	misc_feature	/*tag= h	
FT		/note= "3", near consensus site"	
FT		4798..4817	
FT	misc_feature	/*tag= i	
FT		/note= "3", near consensus site"	
FT		5023..5045	
FT	misc_feature	/*tag= j	
FT		/note= "3", near consensus site"	
FT		5333..5355	
FT	misc_feature	/*tag= k	
FT		/note= "3", near consensus site"	
FT		5604..5632	
FT	misc_feature	/*tag= m	
FT		/note= "3", near consensus site"	
FT		5717..5745	
FT	misc_feature	/*tag= n	
FT		/note= "3", near consensus site"	
FT		6239..6258	
FT	misc_feature	/*tag= o	
FT		/note= "3", near consensus site"	
FT		6658..6682	
FT	misc_feature	/*tag= p	
FT		/note= "3", near consensus site"	
FT		7159..7176	
FT	misc_feature		

Query Match	Best Local Similarity	Matches 4420: Conservative	Score 55.5%	Pred. No. 0:	Mismatches 14:	Indels 0:	Gaps 0:
415	ccacatgcgaataagcctccacacgtctcttcttgcgtcttgcacatctgctta	474					
2960	ccacatcgtttatagctctccacacgtcttcttcttgcgtcttgcacatctgctta	3019					
475	gtgcacacgaagaatactacctggtgtgcagtgagtgaaactgcatacgtggactatgcacaagt	534					
3020	gtgcacacgaagaatactacctggtgtgcagtgagtgaaactgcatacgtggactatgcacaagt	3079					



QY 535 atctcgtgtagctgctctgttgagacgaagatcttcctccttagagtgccaaaatctttccat 594  
|||||  
Db 3080 atctcgtgtagctgctctgttgagacgaagatcttcctccttagagtgccaaaatctttccat 3139  
QY 595 tcaacaactcagctcggtgtacaanaaagactcgtctgttagaattcaacggttcacctttca 654  
|||||  
Db 3140 tcaacaactcagctcggtgtacaanaaagactctgttagaattcaacggttcacctttca 3199  
QY 655 acatcgcgaacgaagcgaaccccttgatgggtctgttaggtcttcctccatccagcgcggaag 714  
|||||  
Db 3200 acatcgcgaacgaagcgaaccccttgatgggtctgttaggtcttcctccatccagcgcggaag 3259  
QY 715 ttatagatcacgtgtgtcatctaacacttaagaacatgtctcccatcctgtcaactcttcag 774  
|||||  
Db 3260 ttatagatcacgtgtgtcatctaacacttaagaacatgtctcccatcctgtcaactcttcag 3319  
QY 775 ctgtctgtgtatcctactcgtgaagaagcttctgagggagctgaaatagatgatacgaacagtc 834  
|||||  
Db 3320 ctgtctgtgtatcctactcgtgaagaagcttctgagggagctgaaatagatgatacgaacagtc 3379  
QY 835 aaagggaagaagaagatataaagctcttcctcctggtggaagccatataatgtctggaag 894  
|||||  
Db 3380 aaagggaagaagaagatataaagctcttcctcctggtggaagccatataatgtctggaag 3439  
QY 895 tctctgaagagaaatggttcacaaatggtccttgaaacccactgtgaccttaacctatattctt 954  
|||||  
Db 3440 tctctgaagagaaatggttcacaaatggtccttgaaacccactgtgaccttaacctatattctt 3499  
QY 955 ctcatctgtgacctgtgttaaaagaacttgaaattcagagccctattgtgagccctactagtatgta 1014  
|||||  
Db 3500 ctcatctgtgacctgtgttaaaagaacttgaaattcagagccctattgtgagccctactagtatgta 3559  
QY 1015 gagaagggagctctgagcccaagaaagacacagaccttgacaaaattatatactatttg 1074  
|||||  
Db 3560 gagaagggagctctgagcccaagaaagacacagaccttgacaaaattatatactatttg 3619  
QY 1075 ctgtatttgatgaaagggaagaagttggtcacctcagaacaagaactccttgatgtagcagata 1134  
|||||  
Db 3620 ctgtatttgatgaaagggaagaagttggtcacctcagaacaagaactccttgatgtagcagata 3679  
QY 1135 gggatgtcgtcatctgtctgggctgtgagcctaataatgacacagtcgaatgtgtataataa 1194  
|||||  
Db 3680 gggatgtcgtcatctgtctgggctgtgagcctaataatgacacagtcgaatgtgtataataa 3739  
QY 1195 ggtctctgcaggtctgtgattgtagtgcacaggaataacagtcattgtggcctgtgattgga 1254  
|||||  
Db 3740 ggtctctgcaggtctgtgattgtagtgcacaggaataacagtcattgtggcctgtgattgga 3799  
QY 1255 tgggacacacatctctgaaatgtagctatcatatctctcgaaggtcacacatcttctgtgagga 1314  
|||||  
Db 3800 tgggacacacatctctgaaatgtagctatcatatctctcgaaggtcacacatcttctgtgagga 3859  
QY 1315 accatctgcaggggtctctctgtgaaatctctgcacataacttctcttactgtctcaaaacct 1374  
|||||  
Db 3860 accatctgcaggggtctctctgtgaaatctctgcacataacttctcttactgtctcaaaacct 3919  
QY 1375 tgaatggacctgtgagcagttcttactggtttgttcataatctctcccaacaatgtagca 1434  
|||||  
Db 3920 tgaatggacctgtgagcagttcttactggtttgttcataatctctcccaacaatgtagca 3979  
QY 1435 tgggaagcttatgtcaaaatgtagcagctgttccagaggaaccccaactacgaatgtaaaaaa 1494  
|||||  
Db 3980 tgggaagcttatgtcaaaatgtagcagctgttccagaggaaccccaactacgaatgtaaaaaa 4039  
QY 1495 atggaagaagcgggaagacatgatatgatatgtactagattcttgaagaatgtagtggtcaggt 1554  
|||||  
Db 4040 atggaagaagcgggaagacatgatatgatatgtactagattcttgaagaatgtagtggtcaggt 4099  
QY 1555 ttgagtgtgaacaactctctcctcttataccaatctgcctcagtttcgaagaagcattccta 1614  
|||||  
Db 4100 ttgagtgtgaacaactctctcctcttataccaatctgcctcagtttcgaagaagcattccta 4159

QY 1615 aaacttggtatcatcatatgtctgtctgaaaggaaggaactggtgactatgtctccttagtc 1674  
|||||  
Db 4160 aaacttggtatcatcatatgtctgtctgaaaggaaggaactggtgactatgtctccttagtc 4219  
QY 1675 tggcccccgtatgacagaagttataaagtccaatatcttgaacaatggtccctcagcgatttg 1734  
|||||  
Db 4220 tggcccccgtatgacagaagttataaagtccaatatcttgaacaatggtccctcagcgatttg 4279  
QY 1735 gtagaagtgtacaaaaaagttccgaatttatgtgcatcaacagatgaaacctttaagactgtg 1794  
|||||  
Db 4280 gtagaagtgtacaaaaaagttccgaatttatgtgcatcaacagatgaaacctttaagactgtg 4339  
QY 1795 aagcttatcagatgataatcagaagatcttgggaaccttacttatgtgggaagttggaaga 1854  
|||||  
Db 4340 aagcttatcagatgataatcagaagatcttgggaaccttacttatgtgggaagttggaaga 4399  
QY 1855 cactgtgtatatatttaagaatcaagaagcagacatatatacctataccttcacggaa 1914  
|||||  
Db 4400 cactgtgtatatatttaagaatcaagaagcagacatatatacctataccttcacggaa 4459  
QY 1915 tcaactggtccgtctcttctgttatcaagaagatataccaaaaggtgtiaaaacatttgaag 1974  
|||||  
Db 4460 tcaactggtccgtctcttctgttatcaagaagatataccaaaaggtgtiaaaacatttgaag 4519  
QY 1975 atttccaattctgcaggaagaataatcaataataatgagcagtgactgtagaagatg 2034  
|||||  
Db 4520 atttccaattctgcaggaagaataatcaataataatgagcagtgactgtagaagatg 4579  
QY 2035 ggtccacataatcgaatctctcgtgtgcctgagaccgctattactagtttcgttaatttg 2094  
|||||  
Db 4580 ggtccacataatcgaatctctcgtgtgcctgagaccgctattactagtttcgttaatttg 4639  
QY 2095 aagagatcttagcttcaagactcattgtgacctctctcctatctgtctacaagaactcttag 2154  
|||||  
Db 4640 aagagatcttagcttcaagactcattgtgacctctctcctatctgtctacaagaactcttag 4699  
QY 2155 atcaaaagsgnaaaccaagataatgttcagacaagaaggaatgtcatccgtttctgtatttg 2214  
|||||  
Db 4700 atcaaaagsgnaaaccaagataatgttcagacaagaaggaatgtcatccgtttctgtatttg 4759  
QY 2215 atgagaaccgaagctgtgtacccctcagaggaatatacaacggtttctcccaatccagct 2274  
|||||  
Db 4760 atgagaaccgaagctgtgtacccctcagaggaatatacaacggtttctcccaatccagct 4819  
QY 2275 gagtgcaacttgaaagatccagagttcccaagctccaacataltgacaagcaltatgct 2334  
|||||  
Db 4820 gagtgcaacttgaaagatccagagttcccaagctccaacataltgacaagcaltatgct 4879  
QY 2335 atgtttttgatagttttgagttgtcagttgtgttgcatggaaggtgcatactgtatactc 2394  
|||||  
Db 4880 atgtttttgatagttttgagttgtcagttgtgttgcatggaaggtgcatactgtatactc 4939  
QY 2395 taagatgggaagcaagaactgactctcttctgtctctctctcctcgtgatatataccttcaac 2454  
|||||  
Db 4940 taagatgggaagcaagaactgactctcttctgtctctctctcctcgtgatatataccttcaac 4999  
QY 2455 acaaaatggtctatgaaacaacacactacacactatcccatcttcagaggaactgtctca 2514  
|||||  
Db 5000 acaaaatggtctatgaaacaacacactacacactatcccatcttcagaggaactgtctca 5059  
QY 2515 tgtcgaatggaaaacccaggtctatagatctctgggtgtgccaaactcagacttcggaaca 2574  
|||||  
Db 5060 tgtcgaatggaaaacccaggtctatagatctctgggtgtgccaaactcagacttcggaaca 5119  
QY 2575 gaggcataccgcttactggaaggtttcttagtttgacaagaacactggtgattatagg 2634  
|||||  
Db 5120 gaggcataccgcttactggaaggtttcttagtttgacaagaacactggtgattatagg 5179  
QY 2635 aggaagttatgaaagataattccaacatactgtcgtgaatlaaanaacatgtcaatgaaacaa 2694  
|||||  
Db 5180 aggaagttatgaaagataattccaacatactgtcgtgaatlaaanaacatgtcaatgaaacaa 5239  
QY 2695 gaagcttctccagaattcaagaacaccttagcactaggcaaaagaaatttaatgtcaccc 2754



|||||  
Db 5240 gaagctctccagagattcaagacaccctagacagtaagcaaatltaatgccacc 5239  
Oy 2755 caccagctcttgaagagccatcaagcggaataaactcgtactactcttcagtcagatcaag 2814  
Db 5300 caccagctcttgaagagccatcaagcggaataaactcgtactactcttcagtcagatcaag 5359  
Oy 2815 aggaatctgactatgatataccatcatcagttgaaatgaagaagaatttgaactt 2874  
Db 5360 aggaatctgactatgatataccatcatcagttgaaatgaagaagaatttgaactt 5419  
Oy 2875 atgataagatgaaataatcaagcccccgcaagcttccaagaanaaacagacatacttca 2934  
Db 5420 atgataagatgaaataatcaagcccccgcaagcttccaagaanaaacagacatacttca 5479  
Oy 2935 ttgctgcagttgagaagcgtctcggaattatggatgagtagctcccaatgttctaaga 2994  
Db 5480 ttgctgcagttgagaagcgtctcggaattatggatgagtagctcccaatgttctaaga 5539  
Oy 2995 acaaggctcagaatgagcagttccctcagttccaagaagaattgtttccaagaatttactg 3054  
Db 5540 acaaggctcagaatgagcagttccctcagttccaagaagaattgtttccaagaatttactg 5599  
Oy 3055 atggctccttactcaagccttataccgttggagaactaaatgaaacatttggactccg 3114  
Db 5600 atggctccttactcaagccttataccgttggagaactaaatgaaacatttggactccg 5659  
Oy 3115 ggcacatataagcagcagaattgaagaataatcatcagtaacttcagaatacaagcct 3174  
Db 5660 ggcacatataagcagcagaattgaagaataatcatcagtaacttcagaatacaagcct 5719  
Oy 3175 ctgctccctatctcttctatctagccttattcttataagaaagatcaagcgcaagag 3234  
Db 5720 ctgctccctatctcttctatctagccttattcttataagaaagatcaagcgcaagag 5779  
Oy 3235 cagaacactagaanaaacttctcaagcctaaatgaacccaacttacttctggaaatgctc 3294  
Db 5780 cagaacactagaanaaacttctcaagcctaaatgaacccaacttacttctggaaatgctc 5839  
Oy 3295 aacatcataatgacccactaaagaatgagttgactgtaagcctggcttacttctctg 3354  
Db 5840 aacatcataatgacccactaaagaatgagttgactgtaagcctggcttacttctctg 5899  
Oy 3355 atgttgacctggaanaaagatgtgcactcagcgccgattggaacccctctgtctgcaca 3414  
Db 5900 atgttgacctggaanaaagatgtgcactcagcgccgattggaacccctctgtctgcaca 5959  
Oy 3415 ctaaacacatgaaacctgtcatctggagacaagtgcacagtaacaggaatttgcctgttt 3474  
Db 5960 ctaaacacatgaaacctgtcatctggagacaagtgcacagtaacaggaatttgcctgttt 6019  
Oy 3475 tcaacatcttctgagagacccaagaagctgtactccaactgaaataatgaaagaactgtca 3534  
Db 6020 tcaacatcttctgagagacccaagaagctgtactccaactgaaataatgaaagaactgtca 6079  
Oy 3535 gggctccctgcaataatcagaatggaagatcccaactttaagaagaatatatgccttccag 3594  
Db 6080 gggctccctgcaataatcagaatggaagatcccaactttaagaagaatatatgccttccag 6139  
Oy 3595 caatcaatgctacataatgatacatcactcgtttaglaatgtgctcagagtaacaagga 3654  
Db 6140 caatcaatgctacataatgatacatcactcgtttaglaatgtgctcagagtaacaagga 6199  
Oy 3655 ttcgatgtatctgtcagcagctggcagcaatgaaatacccatctatctatcttcaagt 3714  
Db 6200 ttcgatgtatctgtcagcagctggcagcaatgaaatacccatctatctatcttcaagt 6259  
Oy 3715 gacatgttctaacgttaagaaaaaagagagtaaaatggcactgtacaaatctcatac 3774  
Db 6260 gacatgttctaacgttaagaaaaaagagagtaaaatggcactgtacaaatctcatac 6319  
Oy 3775 caagtgctttttgagacagttgaaatgttaccatccaagaagctggaatttggcgggtggaat 3834  
|||||

Db 6320 caagtgctttttgagacagttgaaatgttaccatccaagaagctggaatttggcgggtggaat 6379  
Oy 3835 gcttatgtggcagcatctatcatgtctggatgaacacacttttctgtgtacagcata 3884  
Db 6380 gcttatgtggcagcatctatcatgtctggatgaacacacttttctgtgtacagcata 6439  
Oy 3895 agttcagaacctcccttggaaatgtctcttgacaacatataagatttccaatctaacgtt 3954  
Db 6440 agttcagaacctcccttggaaatgtctcttgacaacatataagatttccaatctaacgtt 6499  
Oy 3955 caggacaataatgaaatgagcccccaagctgagcaactcatatctccgataatca 4014  
Db 6500 caggacaataatgaaatgagcccccaagctgagcaactcatatctccgataatca 6559  
Oy 4015 atgcttgagacacaaagagcccttcttctgatacaagtgatctgttgcaccaatga 4074  
Db 6560 atgcttgagacacaaagagcccttcttctgatacaagtgatctgttgcaccaatga 6619  
Oy 4075 ttatccaagcatcaagacccaggtgcccgtccaagaattctccaagcctctacatctc 4134  
Db 6620 ttatccaagcatcaagacccaggtgcccgtccaagaattctccaagcctctacatctc 6679  
Oy 4135 agttatcatcatgtatagttctgataaggaagttgagacattatccgaagaaatcca 4194  
Db 6680 agttatcatcatgtatagttctgataaggaagttgagacattatccgaagaaatcca 6739  
Oy 4195 ctggaaaccttaatgtctcttcttgcaatgttgatctatctcggagataaaacacataat 4254  
Db 6740 ctggaaaccttaatgtctcttcttgcaatgttgatctatctcggagataaaacacataat 6799  
Oy 4255 ttaacccccaatattatgctcgatacatccgttggacccaactcatatagatctgca 4314  
Db 6800 ttaacccccaatattatgctcgatacatccgttggacccaactcatatagatctgca 6859  
Oy 4315 gcaactctgcataagatgtgagctgtgatatcaaatagttgacagcatgcaattggaa 4374  
Db 6860 gcaactctgcataagatgtgagctgtgatatcaaatagttgacagcatgcaattggaa 6919  
Oy 4375 tggagagtaagaagcatatcagaatgacagaattactgtcttaacttactccaatatgt 4434  
Db 6920 tggagagtaagaagcatatcagaatgacagaattactgtcttaacttactccaatatgt 6979  
Oy 4435 ttggcaacctgtctcctctcaaaaagctcgactccaactccaaggagagtaatctctgca 4494  
Db 6980 ttggcaacctgtctcctctcaaaaagctcgactccaactccaaggagagtaatctctgca 7039  
Oy 4495 gacctcaggtgaataatccaagaagttgctgcaagtggacttccagaagaacatgaag 4554  
Db 7040 gacctcaggtgaataatccaagaagttgctgcaagtggacttccagaagaacatgaag 7099  
Oy 4555 tcaacagagtaactacatccaaggaataaattctcgtctaccagcatgtatgtgaagaagt 4614  
Db 7100 tcaacagagtaactacatccaaggaataaattctcgtctaccagcatgtatgtgaagaagt 7159  
Oy 4615 tcccatctccagcagtaaaatgagcaatgagtgactctcttttccaagaatggcaag 4674  
Db 7160 tcccatctccagcagtaaaatgagcaatgagtgactctcttttccaagaatggcaag 7219  
Oy 4675 taaaggttttcaaggaatacaagactcctcaacacgtgtgtaactctctlagaccac 4734  
Db 7220 taaaggttttcaaggaatacaagactcctcaacacgtgtgtaactctctlagaccac 7279  
Oy 4735 cgttactgactcgtctactcttccaatccaaccccaagattgggtgacacagattgcctga 4794  
Db 7280 cgttactgactcgtctactcttccaatccaaccccaagattgggtgacacagattgcctga 7339  
Oy 4795 ggaatggaggtctggtcgtcagagcacaggaacctctcgtcagagcaggtc 4848  
Db 7340 ggaatggaggtctggtcgtcagagcacaggaacctctcgtcagagcaggtc 7393  
|||||

RESULT 6  
AAQ76016



ID AA076016 standard; cDNA: 4629 BP.  
XX  
AC AA076016;  
XX  
DT 20-JUL-1995 (first entry)  
XX  
DE B-domain deleted Factor-VIII.  
XX  
XX Factor-VIII; blood-clotting; hemophilia A; gene therapy;  
KM adenovirus; vector; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO9429471-A.  
XX  
PD 22-DEC-1994.  
XX  
PF 13-APR-1994; 94MO-US04075.  
XX  
PR 10-JUN-1993; 93US-0074920.  
PR 25-MAR-1994; 94US-0218335.  
XX  
PA (GENE-) GENETIC THERAPY INC.  
XX  
PI Connelly S, Kaleko M, Smith T;  
XX  
DR WPI; 1995-036495/05.  
XX  
DR P-PSDB; AAR67709.  
PT New adenoviral vectors for treatment of haemophilia - contg. a  
PT DNA sequence encoding a clotting factor, partic. Factor VIII or  
PT Factor IX  
XX  
PS Disclosure; Fig. 17A-17C; 116pp; English.  
XX  
CC Human Factor-VIII cDNA, from which the B domain had been deleted, was  
CC used to construct recombinant adenovirus vectors that produced  
CC therapeutic levels of the clotting factor when administered to an  
CC animal host, potentially providing hemophilia A gene therapy.  
XX  
XX  
SO Sequence 4629 BP; 1318 A; 1051 C; 1018 G; 1242 T; 0 other;

Query Match 54.8%; Score 4355.4; DB 16; Length 4629;  
Best Local Similarity 99.0%; Pred. No. 0;  
Matches 4383; Conservatlve 0; Mismatches 46; Indels 0; Gaps 0;

QY 420 atgcaaatagagctcttcacactgctcttcttctgctcttcttgagattcgtcttagtgc 479  
DB 1 atgcaaatagagctcttcacactgctcttcttctgctcttcttgagattcgtcttagtgc 60  
QY 480 accgaagaatactacctggtgcaatggaactgcatggaactatagcaaatgcatcc 539  
DB 61 accgaagaatactacctggtgcaatggaactgcatggaactatagcaaatgcatcc 120  
QY 540 ggtgagctgctgtggaagcaagattccctcctagagtgcaaaatctttccattcaac 599  
DB 121 ggtgagctgctgtggaagcaagattccctcctagagtgcaaaatctttccattcaac 180  
QY 600 acccgaatcggtatcaaaaagactctgtttgtagaatccaaggttccactttcaaac 659  
DB 181 acccgaatcggtatcaaaaagactctgtttgtagaatccaaggttccactttcaaac 240  
QY 660 gctaagccaagccaaccttgatggtctgtctaggtccacacacacaggttattat 719  
DB 241 gctaagccaagccaaccttgatggtctgtctaggtccacacacacaggttattat 300  
QY 720 gatacagtgtgcatatcaacttaagaacaaggcttccacacacacaggttattat 779  
DB 301 gatacagtgtgcatatcaacttaagaacaaggcttccacacacacaggttattat 360  
QY 780 ggtgtatctactatgaaagctcttgaggagctgaatatgatgatcagacagctaaag 839  
|||||

DB 361 ggtgtatctactatgaaagctcttgaggagctgaatatgatgatcagacagctaaag 420  
QY 840 gagaagaagaatgataaagctcttccctggtggaagccaataatgtctggaagctctg 899  
DB 421 gagaagaagaatgataaagctcttccctggtggaagccaataatgtctggaagctctg 480  
QY 900 aagaagaatgtlccaatgtgctctgacccactgtgcttactactacatattcttccat 959  
DB 481 aagaagaatgtlccaatgtgctctgacccactgtgcttactactacatattcttccat 540  
QY 960 gtgagactgttaaaagacttgaattcaggtccatattggaagccactactagtatgaagaa 1019  
DB 541 gtgagactgttaaaagacttgaattcaggtccatattggaagccactactagtatgaagaa 600  
QY 1020 gggagcttgcccaagaaagacacagacactgtgcaacaattactactcttggcgta 1079  
DB 601 gggagcttgcccaagaaagacacagacactgtgcaacaattactactcttggcgta 660  
QY 1080 ttgtatgaaggggaaaagtgtgcaatcagaacaagaaaccccttgatgcaagataggat 1139  
DB 661 ttgtatgaaggggaaaagtgtgcaatcagaacaagaaaccccttgatgcaagataggat 720  
QY 1140 gctgcatctgctcggtcggtcggttaaatgcaacaagtcattgttattgttaacaggtct 1199  
DB 721 gctgcatctgctcggtcggtcggttaaatgcaacaagtcattgttattgttaacaggtct 780  
QY 1200 ctgcaagctgtgattggaatgcccaggaatcagttcttggatgattggaatggc 1259  
DB 781 ctgcaagctgtgattggaatgcccaggaatcagttcttggatgattggaatggc 840  
QY 1260 accactctgaaagtgcactcaatatctcctgaaggtcacacattctgttgaagaacct 1319  
DB 841 accactctgaaagtgcactcaatatctcctgaaggtcacacattctgttgaagaacct 900  
QY 1320 cgcagcggtctcttgaatctcgcaaatcttcccttactgttcaacaactcttgatg 1379  
DB 901 cgcagcggtctcttgaatctcgcaaatcttcccttactgttcaacaactcttgatg 960  
QY 1380 gaccttggaagcttctctactgtttgtcatatctcttcccaacactgagatgaa 1439  
DB 961 gaccttggaagcttctctactgtttgtcatatctcttcccaacactgagatgaa 1020  
QY 1440 gcttatgtcaaatgtagacagctgttccagaagaaacccaactagaaatgaataatgaa 1499  
DB 1021 gcttatgtcaaatgtagacagctgttccagaagaaacccaactagaaatgaataatgaa 1080  
QY 1500 gaagcggaagactagatgattacttacttacttgaatgggtgtggtcaggtttgat 1559  
DB 1081 gaagcggaagactagatgattacttacttacttgaatgggtgtggtcaggtttgat 1140  
QY 1560 gatgaacactctctcttcttataccaatctgctcagttgccaagaagcatcctaaact 1619  
DB 1141 gatgaacactctctcttcttataccaatctgctcagttgccaagaagcatcctaaact 1200  
QY 1620 tgggtacattacattgctgtctgaaagaggagactgtgactatgctccctagtcctgc 1679  
DB 1201 tgggtacattacattgctgtctgaaagaggagactgtgactatgctccctagtcctgc 1260  
QY 1680 cccgaatgacagaagtataaaagattgaatattgaaatggccttcagcggaattgttag 1739  
DB 1261 cccgaatgacagaagtataaaagattgaatattgaaatggccttcagcggaattgttag 1320  
QY 1740 aagtaaaaaaagtcggaattatggtcatacaatgaacacttgaagactcggtgaagct 1799  
DB 1321 aagtaaaaaaagtcggaattatggtcatacaatgaacacttgaagactcggtgaagct 1380  
QY 1800 attcagcatgaaatcagaatcttgggaacttacttatttgggaagtgtggaacacactg 1859  
DB 1381 attcagcatgaaatcagaatcttgggaacttacttatttgggaagtgtggaacacactg 1440  
QY 1860 ttgattatatttaagaatcaagcaagcagacacataataactctaccctcaggaataact 1919  
DB 1441 ttgattatatttaagaatcaagcaagcagacacataataactctaccctcaggaataact 1500  
|||||



QY	1920	gagttccgtcccttgttatctcaagagataccaagaagtgtaaaacatctgaagatttc	1979
Db	1501	gatttccgtcccttgttatctcaagagataccaagaagtgtaaaacatctgaagatttc	1560
QY	1980	ccaattctgcacagagaataattccaataataatgacagtgactgtgaagatggcca	2039
Db	1561	ccaattctgcacagagaataattccaataataatgacagtgactgtgaagatggcca	1620
QY	2040	actaatcagatccctcgtgtcgtaccgcgtatctactagtttcgttaataatgagaga	2099
Db	1621	actaatcagatccctcgtgtcgtaccgccttactactagtttcgttaataatgagaga	1680
QY	2100	gattctagtttcagagactcagatctgtgcctctccctccatctgtctcaagaagatctgtatcca	2159
Db	1661	gattctagtttcagagactcagatctgtgcctctccctccatctgtctcaagaagatctgtatcca	1740
QY	2160	agagaaaccaagaataatgtlcaagaacaagagaaatgtcatccctgttctcgtatctgtatgag	2219
Db	1741	agagaaaccaagaataatgtlcaagaacaagagaaatgtcatccctgttctcgtatctgtatgag	1800
QY	2220	aaccgaagctgtgtactctcacagagaatacaacgcttcttcccaatccagctgtgagtg	2279
Db	1801	aaccgaagctgtgtactctcacagagaatacaacgcttcttcccaatccagctgtgagtg	1860
QY	2280	cagcttgagagatccagaagttccagaagcttcccaacatcatgacgaagaatcatatggttatggt	2339
Db	1861	cagcttgagagatccagaagttccagaagcttcccaacatcatgacgaagaatcatatggttatggt	1920
QY	2340	ttgtatagttcttgcaatgtgtcaagtttctgtgcatagtgagtcgaactgttatcatctgaagc	2399
Db	1921	ttgtatagttcttgcaatgtgtgttctgtgcataagctcgtgactgtgtatcatctgaagc	1980
QY	2400	attggaacacagaagctgaactctcttctgtcctctctctctcgtatataactcttaaacacaaa	2459
Db	1961	attggaacacagaagctgtactctctctctctctctctctctcgtatataactcttaaacacaaa	2040
QY	2460	attgtctcattgaagaacacacatccacatcccatctccatctccagaagaacatgtcttatctgcg	2519
Db	2041	attgtctcattgaagaacacacatccatctccatctccagaagaacatgtcttatctgcg	2100
QY	2520	attgaaaaaccagtgcttatgtatgtctcgtgggtgtccacaactcagactcttcggaacaagagc	2579
Db	2101	attgaaaaaccagtgcttatgtatgtctcgtgggtgtccacaactcagactcttcggaacaagagc	2160
QY	2580	attgacgcgccttactcgaaggttctctagtctgtgaacaagaacatcgtgtatctatcgaagagac	2639
Db	2161	attgacgcgccttactcgaaggttctctagtctgtgaacaagaacatcgtgtatctatcgaagagac	2220
QY	2640	agttatgagaatattctcagcatctgtcgtgagtaaaaaaatgycatctgaaaccaagaagc	2699
Db	2221	agttatgagaatattctcagcatctgtcgtgagtaaaaaaatgycatctgaaaccaagaagc	2280
QY	2700	ttctcccaagaattccaagaaccccttagcactatgagcaaaagcaatttaatgcccccccaca	2759
Db	2261	ttctcccaagaattccaagaaccccttagcactatgagcaaaagcaatttaatgcccccccaca	2340
QY	2760	gtctctgaaacgcatacaacggaagaataactctgtactactcttccatctcagatccaagaagaa	2819
Db	2341	gtctctgaaacgcatacaacggaagaataactctgtactactcttccatctcagatccaagaagaa	2400
QY	2820	attgactatgatatatccatatcatgtttgaaatggaagaaggaagattcttgacaatttatgat	2879
Db	2401	attgactatgatatatccatatcatgtttgaaatggaagaaggaagattcttgacaatttatgat	2460
QY	2880	gaggaatgaataacataagccccgcagcttccaagaagaacaacgcacatatttatgtc	2939
Db	2461	gaggaatgaataacataagccccgcagcttccaagaagaacaacgcacatatttatgtc	2520
QY	2940	gcagctgagaagctcttgggaattatgtgagtgatgctccccaacatglttctaagaacag	2999
Db	2521	gcagctgagaagctcttgggaattatgtgagtgatgctccccaacatglttctaagaagag	2580

QY	3000	gqccaagatgagcagtgctccctcagttccaaagaagttgcttccaggaaattacatgagc	3059
Db	2581	gctcagagttggcagtgctccctcagttccaaagaagttgcttccaggaaattacatgagc	2640
QY	3060	tcctttactcagcccttaacccgltggagaactcaaaacatttggagctccctgggcca	3119
QY	3120	tataaagaacagaagtttgaagaataataatgtaactttcagaataacagccctcgt	3179
Db	2701	tataaagaacagaagtttgaagaataataatgtaactttcagaataacagccctcgt	2760
QY	3180	ccctattccctctatctctagccctatttctttagaggaagatccagagcagagca	3239
QY	3240	ccctattccctctatctctagccctatttctttagaggaagatccagagcagagca	3299
Db	2821	ccctagaaaaaaccttctgcaagccctaaagaaacaaacttcttggaaagtgaacat	2880
QY	3300	catagagaccccaataaagaatgagttgagatgcaagccttggcctattctctgatt	3359
Db	2881	catatgagaccccaataaagaatgagttgagatgcaagccttggcctattctctgatt	2940
QY	3360	gaacctggaaaaaagaatgcaactcagagccgatgtagacccctctgctctgcacactaac	3419
Db	2941	gaacctggaaaaaagaatgcaactcagagccgatgtagacccctctgctctgcacactaac	3000
QY	3420	acaactgaacccctgctcatgtggagagaagtgacagatccaggaattgctctgttttcaac	3479
Db	3001	acaactgaacccctgctcatgtggagagaagtgacagatccaggaattgctctgctgcagagct	3060
QY	3480	atctttgattggaccaaaagctctgtaacttcacatgaaataatgaaagaacatgcagagct	3539
Db	3061	atctttgattggaccaaaagctctgtaacttcacatgaaataatgaaagaacatgcagagct	3120
QY	3540	ccctgcacatatccagatgtagaagaatcccaacttttaagaagaaattatcgtctccatgcacac	3599
Db	3121	ccctgcacatatccagatgtagaagaatcccaacttttaagaagaaattatcgtctccatgcacac	3180
QY	3600	aatggtcacataaatvggataacatacctgctgttaagttaatggtcccaagatccaaagatttga	3659
Db	3181	aatggtcacataaatvggataacatacctgctgttaagttaatggtcccaagatccaaagatttga	3240
QY	3660	tgttatctgcctcagcagatgtaggcaagaatgaaacatccattccattccattccatgagat	3719
Db	3241	tgttatctgcctcagcagatgtaggcaagaatgaaacatccattccattccattccatgagat	3300
QY	3720	gtgttcaactgtacgaaaaaaagagagtaaaatggtgcagtacacatctctatccaggt	3779
Db	3301	gtgttcaactgtacgaaaaaaagagagtagaaatggtgcagtacacatctctatccaggt	3360
QY	3780	gtttttgagacagtgagaaatggttacatccacaaagctcgaaatttggcgggtvggaatgctt	3839
Db	3361	gtttttgagacagtgagaaatggttacatccacaaagctcgaaatttggcgggtvggaatgctt	3420
QY	3840	attggcagagcactacaatgtctggatgtagcacaacttttctgtgttacaagcaataaggt	3899
Db	3421	attggcagagcactacaatgtctggatgtagcacaacttttctgtgttacaagcaataaggt	3480
QY	3900	cagagctcccccgggaaatgagcttctctggacaacatagaagatttccagattacagcttcaga	3959
Db	3481	cagagctcccccgggaaatgagcttctctggacaacatagaagatttccagattacagcttcaga	3540
QY	3960	caataatgagacagtggtgcccaaaagctgtgccagagcttcatattccggataatcaatggtc	4019
Db	3541	caataatgagacagtggtgcccaaaagctgtgccagagcttcatattccggataatcaatggtc	3600
QY	4020	tggagacccaaagggcccttttcttggatcaagaatgtagatctgttggacccaatattatt	4079
Db	3601	tggagacccaaagggcccttttcttggatcaagaatgtagatctgttggacccaatattatt	3660
QY	4080	cacggatccaagaccccaaggttgcccgctcaggaagttccacagcccttaacatcttccagtt	4139



Db	3661	caacgcatcaagaccgaaggtgcccgtcaagaatctccagccctctacatctccagttt	3720
Qy	4140	atcatcgtatgatctctgtatgggaagaagtggcagaacttaicgaggaataatctcattga	4199
Db	3721	atcatcagtatagatctctgtatgggaagaagtggcagaacttaicgaggaataatctcattga	3780
Qy	4200	accttaatggtctctctttgtgcaatgtygatcatctcyggataaacaacaatatcttaac	4259
Db	3781	accttaatggtctctctttgtgcaatgtygatcatctcyggataaacaacaatatcttaac	3840
Qy	4260	ccctcaatatgtctctgatcatcccggtttgaacccaactcaatrlaagatctgcagacac	4319
Db	3841	ccctcaatatgtctctgatcatcccggtttgaacccaactcaatrlaagatctgcagacac	3900
Qy	4320	ctctcgatgtgagttctgaatggcgttgatcttaataatgtctgcagcatctgtggaatggag	4379
Db	3901	ctctcgatgtgagttctgaatggcgttgatcttaataatgtctgcagcatctgtggaatggag	3960
Qy	4380	agtaaaagcaatatacagatgcacagatactgcttcaactcttaacttaccataatgttcgc	4439
Db	3961	agtaaaagcaatatacagatgcacagatactgcttcaactcttaacttaccataatgttcgc	4020
Qy	4440	acctgtctcccttcaaaaagcttgactcaactcccaaggagagtaatgtcctggaagcct	4499
Db	4021	acctgtctcccttcaaaaagcttgactcaactcccaaggagagtaatgtcctggaagcct	4080
Qy	4500	caagtgtaataatcacaagaagtgctgtcgaagtgtgacttccgaaagacaatgtaaagtca	4559
Db	4081	caagtgtaataatcacaagaagtgctgtcgaagtgtgacttccgaaagacaatgtaaagtca	4140
Qy	4560	ggaagtaactactcgaaggagtaaaatctctgtcttaacagatgatagtgaaagattccctc	4619
Db	4141	ggaagtaactactcgaaggagtaaaatctctgtcttaacagatgatagtgaaagattccctc	4200
Qy	4620	acctcagcagtgtaagaatgtgcatactgagttactcttttttcgaatgtgaaagttaag	4679
Db	4201	acctcagcagtgtaagaatgtgcatactgagttactcttttttcgaatgtgaaagttaag	4260
Qy	4680	gttttccaggaatcaagaactccttcaacactgtgtgtaactctctagaccacacgtta	4739
Db	4261	gttttccaggaatcaagaactccttcaacactgtgtgtaactctctagaccacacgtta	4320
Qy	4740	ctgactcgtactactcgaattcacacccccagagttgggtgcacagatgtccctgagatg	4799
Db	4321	ctgactcgtactactcgaattcacacccccagagttgggtgcacagatgtccctgagatg	4380
Qy	4800	gaggtctctgggtctgcgagggcacaggaactctactgactgcagcgagttc	4848
Db	4381	gaggtctctgggtctgcgagggcacaggaactctactgaggtgtgcacatgc	4429
RESULT 7			
AAV23339			
ID	AAV23339	standard; DNM: 4670 BP.	
XX	AAV23339;		
XX	17-AUG-1998	(first entry)	
DE	Human Factor-VIII gene lacking central B domain.		
XX			
XX	Factor-VIII; blood clotting; human; synthetic gene; codon usage; ss.		
OS	Homo sapiens.		
XX	Synthetic.		
PN	W09812207-A1.		
XX			
PD	26-MAR-1998.		
XX			
PF	18-SEP-1997;	97WO-US16639.	
XX			

Query Match	54.8%	Score 4351.2	DB 19	Length 4670
Best Local Similarity	99.7%	Pred No. 0		
Matches 4359	Conservative 0	Mismatches 13	Indels 0	Gaps 0
XX	20-SEP-1996;	96US-0717294.		
XX	(GEHO ) GEN HOSPITAL CORP.			
XX	Haas J, Seed B;			
XX	WPL: 1998-217200/19.			
XX	New synthetic eukaryotic gene(s) - in which non-preferred or less			
XX	preferred codon(s) are replaced to provide high level expression in			
XX	mammalian cell(s)			
PS	Example 3; Fig 12: 92pp; English.			
CC	This gene codes for a human Factor-VIII protein that lacks the			
CC	central B domain (amino acids 760-1639) of the native protein. In			
CC	a novel, claimed synthetic gene (see AAV3288), non-preferred or			
CC	less preferred codons of the native gene are replaced by codons			
CC	favored by highly expressed human genes to provide high-level			
CC	expression in mammalian cells. The synthetic gene was assembled			
CC	from 29 pairs of oligonucleotides (see AAV3340-97) which served as			
CC	PCR templates. Synthetic genes of the invention (see also			
CC	AAV33289-91) are used for production of recombinant proteins in			
CC	mammalian cells at levels of at least 500% of those obtained using			
CC	the natural genes. They can also be used in gene therapy.			
XX				
XX	Sequence 4670 BP; 1325 A; 1072 C; 1036 G; 1237 T; 0 other:			
DB	477 gccaccagaagatactacctggtgtagcagtggaactgcaatggaatataatgcaagtgtat	536		
DB	84 gccaccagaagatactacctggtgtagcagtggaactgcaatggaatataatgcaagtgtat	143		
QY	537 ctcgtgtagctgcctgttggaagcaaatcttcctcctagagtgccaaatctttccatc	596		
DB	144 ctcgtgtagctgcctgttggaagcaaatcttcctcctagagtgccaaatctttccatc	203		
QY	597 aacctcctagctgtgtacaaanaagctcgtttgtagaattcaagtttaacttttcaac	656		
DB	204 aacctcctagctgtgtacaaanaagctcgtttgtagaattcaagtttaacttttcaac	263		
QY	657 atcgcctaaagcaagccaccctcgatgggtctgcctaggtccctcaatccaggctgaagt	716		
DB	264 atcgcctaaagcaagccaccctcgatgggtctgcctaggtccctcaatccaggctgaagt	323		
QY	717 tatgatcacgtgtgtcaattcaacttaagaacatggtctccatcctgttcagcttcattgt	776		
DB	324 tatgatcacgtgtgtcaattcaacttaagaacatggtctccatcctgttcagcttcattgt	383		
QY	777 gttgtgtatctctactacgtgaagctctcttgaggagctggaatataatgaagccagtcaa	836		
DB	384 gttgtgtatctctactacgtgaagctctcttgaggagctggaatataatgaagccagtcaa	443		
QY	837 agggagaagaagatataaagtcttcctcctgttggaagcacaataatgtctcgccaggtc	896		
DB	444 agggagaagaagatataaagtcttcctcctgttggaagcacaataatgtctcgccaggtc	503		
QY	897 ctgaagaagaatggttccaaatggtccttgaccacactgtgtcctactacataatctttct	956		
DB	504 ctgaagaagaatggttccaaatggtccttgaccacactgtgtcctactacataatctttct	563		
QY	957 catgtgtaccctgtgtaaaagactgaaatcgaagctcattatgtgagccctactagtatgtaga	1016		
DB	564 catgtgtaccctgtgtaaaagactgaaatcgaagctcattatgtgagccctactagtatgtaga	623		
QY	1017 gaaggagatcttggccaagaaaagacaacagacttggcacaatatatactacttttgtt	1076		
DB	624 gaaggagatcttggccaagaaaagacaacagacttggcacaatatatactacttttgtt	683		



QY 1077 gtaattgtagaaggaaaagtctggcaactcagaanaaagaagactccttgatcagagatagg 1136  
|||||  
Db 684 gtattgatgaaggaaaagtctggcaactcagaanaaagaagactccttgatcagagatagg 743  
QY 1137 gatctgcattctgtctggcctggcctaaatgcaacagtcgaattgtctgttaaacagg 1196  
|||||  
Db 744 gatctgcattctgtctggcctggcctaaatgcaacagtcgaattgtctgttaaacagg 803  
QY 1197 tctctgcaggcttgatctgattgagccagagaaatcagtcattggctgagatggatagg 1256  
|||||  
Db 804 tctctgcaggcttgatctgattgagccagagaaatcagtcattggctgagatggatagg 863  
QY 1257 ggcacacactcctgaagtgacacccaataatcctcgaaggtcacacaattctctgtgaggaac 1316  
|||||  
Db 864 ggcacacactcctgaagtgacacccaataatcctcgaaggtcacacaattctctgtgaggaac 923  
QY 1317 catcgacagcgctctctggaatctcgccaataaacttcccttaactgctcacaacactcttg 1376  
|||||  
Db 924 catcgacagcgctctctggaatctcgccaataaacttcccttaactgctcacaacactcttg 983  
QY 1377 atggaccttggacagcttctactggtttgtcatalactcttccacccaacatgcatgcatg 1436  
|||||  
Db 984 atggaccttggacagcttctactggtttgtcatalactcttccacccaacatgcatgcatg 1043  
QY 1437 gaagcttatgtlcaaagttagacagctgtccagaggaaacccaactacagaatgaaaataat 1496  
|||||  
Db 1044 gaagcttatgtlcaaagttagacagctgtccagaggaaacccaactacagaatgaaaataat 1103  
QY 1497 gaagaagcggaaagactatgatatgattctactgatatctgtgaaatgtagtgtgcoagttc 1556  
|||||  
Db 1104 gaagaagcggaaagactatgatatgattctactgatatctgtgaaatgtagtgtgcoagttc 1163  
QY 1557 gatgatatacaactctctctcttctatccaatcggctcagtttccagaagaagccctccaaa 1616  
|||||  
Db 1164 gatgatatacaactctctctcttctatccaatcggctcagtttccagaagaagccctccaaa 1223  
QY 1617 acttgggtacattatactgtctgtctgaaaggaggactgggactatgctcccttaagtcctc 1676  
|||||  
Db 1224 acttgggtacattatactgtctgtctgaaaggaggactgggactatgctcccttaagtcctc 1283  
QY 1677 gcccccgtatcacgaagtttataaagtccaattttgaaacatgycctcagcgagatgtgtc 1736  
|||||  
Db 1284 gcccccgtatcacgaagtttataaagtccaattttgaaacatgycctcagcgagatgtgtc 1343  
QY 1737 aggaagatcacaaaagtcgcgaattatggtacatacacagatgaanaacotttaagactcgtgaa 1796  
|||||  
Db 1344 aggaagatcacaaaagtcgcgaattatggtacatacacagatgaanaacotttaagactcgtgaa 1403  
QY 1797 gctattcagatgaatlcagaatccttgggaccttacttatactggggaagcttgagacaca 1856  
|||||  
Db 1404 gctattcagatgaatlcagaatccttgggaccttacttatactggggaagcttgagagacaca 1463  
QY 1857 ctgtttgatattatttaagaatccaagcagagaccataatacatctacacctcaacggaatc 1916  
|||||  
Db 1464 ctgtttgatattatttaagaatccaagcagagaccataatacatctacacctcaacggaatc 1523  
QY 1917 acggaatctccgctccttggtaattcaagagagattaccaaaagctgtaaaaactttggaagat 1976  
|||||  
Db 1524 acggaatctccgctccttggtaattcaagagagattaccaaaagctgtaaaaactttggaagat 1583  
QY 1977 ttccaattctgcagagagaataattcaaatatgaatvggacagtgagctgtagaagatgg 2036  
|||||  
Db 1584 ttccaattctgcagagagaataattcaaatatgaatvggacagtgagctgtagaagatgg 1443  
QY 2037 ccaactaatacagatcctcgtgtgcctggaacccgctatatccttagtctgttataatggag 2096  
|||||  
Db 1644 ccaactaatacagatcctcgtgtgcctggaacccgctatatccttagtctgttataatggag 1703  
QY 2097 agagatctatgcttcaaggaaccatttggccctctcctcaactgctgcatacaagaatctgtagat 2156  
|||||  
Db 1704 agagatctatgcttcaaggaaccatttggccctctcctcaactgctgcatacaagaatctgtagat 1763  
QY 2157 caaagagaaaacagataatgtlcaagacaagagaaigtlcatcctgttttctgtattgat 2216

Db 1764 caaagagaaaacagataatgtlcaagacaagagaaigtlcatcctgttttctgtattgat 1823  
|||||  
QY 2217 gagaacccgaagctcgtgtaccctcacagagaataatacaagcttctccccaatccagcttga 2276  
|||||  
Db 1824 gagaacccgaagctcgtgtaccctcacagagaataatacaagcttctccccaatccagcttga 1883  
QY 2277 gtgcagcttgaagatcccaaggtcccaagcctcccaactcaatgacagatccatgtgctat 2336  
|||||  
Db 1884 gtgcagcttgaagatcccaaggtcccaagcctcccaactcaatgacagatccatgtgctat 1943  
QY 2337 gttttgatagtttgcagttgttcagtttgttgcataaggttgcatactgttatctcta 2396  
|||||  
Db 1944 gttttgatagtttgcagttgttcagtttgttgcataaggttgcatactgttatctcta 2003  
QY 2397 agcatvggacacagacttgactcctctctctgtctctctctctctctctctctctctctct 2456  
|||||  
Db 2004 agcatvggacacagacttgactcctctctctgtctctctctctctctctctctctctctct 2063  
QY 2457 aaaaigtctataagaacacactcaactctccacttccactcagagagaactgtctcatg 2516  
|||||  
Db 2064 aaaaigtctataagaacacactcaactctccacttccactcagagagaactgtctcatg 2123  
QY 2517 tctgattgaaaaacccaggtctcatgattctgtgggtgcacacactcagaacttctcgaaacga 2576  
|||||  
Db 2124 tctgattgaaaaacccaggtctcatgattctgtgggtgcacacactcagaacttctcgaaacga 2183  
QY 2577 ggcatacgcgcttacttgaaagtttcttagttgtgacaagaacactgtgtattctcag 2636  
|||||  
Db 2184 ggcatacgcgcttacttgaaagtttcttagttgtgacaagaacactgtgtattctcag 2243  
QY 2637 gacagttatgaagattttttagcattactgtctgagttaaaaaaatccattgaaaccaaga 2696  
|||||  
Db 2244 gacagttatgaagattttttagcattactgtctgagttaaaaaaatccattgaaaccaaga 2303  
QY 2697 agcttctcccaagaattccaagaaccccttagacactagggcaaaaagaaatttaatgacacca 2756  
|||||  
Db 2304 agcttctcccaagaattccaagaaccccttagacactagggcaaaaagaaatttaatgacacca 2363  
QY 2757 ccaggtcttgaagccgcatccaacgggaaataactcgttactactcttcaagtacataagag 2816  
|||||  
Db 2364 ccaggtcttgaagccgcatccaacgggaaataactcgttactactcttcaagtacataagag 2423  
QY 2817 gaattgactatgatatatacatcatcaagttgaaatgaagaagaattttgacattat 2876  
|||||  
Db 2424 gaattgactatgatatatacatcatcaagttgaaatgaagaagaattttgacattat 2483  
QY 2877 gatgagagatgaaaaatcagaagcccccagcgttccaagaagaaaacacagacatatttatt 2936  
|||||  
Db 2484 gatgagagatgaaaaatcagaagcccccagcgttccaagaagaaaacacagacatatttatt 2543  
QY 2937 gctgcagatggagagagctcgtgggattatgggatagtatagctcccaatgttctaagaaac 2996  
|||||  
Db 2544 gctgcagatggagagagagctcgtgggattatgggatagtatagctcccaatgttctaagaaac 2603  
QY 2997 aggggtccagagtggcaggtgtccctcagttccaagaagaagttgttttccaggaatttactgat 3056  
|||||  
Db 2604 aggggtccagagtggcaggtgtccctcagttccaagaagaagttgttttccaggaatttactgat 2663  
QY 3057 ggtctcttactcagccctataacgltggagaactlaaaltgaacatttgggactcctgggg 3116  
|||||  
Db 2664 ggtctcttactcagccctataacgltggagaactlaaaltgaacatttgggactcctgggg 2723  
QY 3117 ccaatataagacagaagtgtgaagataatataatgtatattctcgaataatcagagcctct 3176  
|||||  
Db 2724 ccaatataagacagaagtgtgaagataatataatgtatattctcgaataatcagagcctct 2783  
QY 3177 cgtccctattccttccattctagccttattcttcttaagaaagatcagaggaagagca 3236  
|||||  
Db 2784 cgtccctattccttccattctagccttattcttcttaagaaagatcagaggaagagca 2843  
QY 3237 gaaactagaaaaaacttgtlcaaagcctlaatgaaaccaaacttacttcttgaaagtgcga 3296  
|||||



Db 2844 gaacctgaaaaaacttctgtaagcctaaatgaaccaaacttacttcttggaagtgcga 2903  
QY 3297 catcatactgaccccaataaagatgagtttgcactgcaagcctgggctattctctgat 3356  
Db 2904 catcatactgaccccaataaagatgagtttgcactgcaagcctgggctattctctgat 2963  
QY 3357 gttgaacctgaaaaaagtctgctactcaagcctgatctgaacccctctgtctgcacact 3416  
Db 2964 gttgaacctgaaaaaagtctgctactcaagcctgatctgaacccctctgtctgcacact 3023  
QY 3417 aaacacatgaaacctgtctgcactgggagacaagtgaacgtaacggaaattgtcctgtcttc 3476  
Db 3024 aaacacatgaaacctgtctgcactgggagacaagtgaacgtaacggaaattgtcctgtcttc 3083  
QY 3477 accatctctgatgagacaaaagctgtgtacttcactgtgaaaaatgtaaaagaacatgcag 3536  
Db 3084 accatctctgatgagacaaaagctgtgtacttcactgtgaaaaatgtaaaagaacatgcag 3143  
QY 3537 gctccctgcacaataatccagaatggaatcccaactttaagaagaattatcgccttcacatgc 3596  
Db 3144 gctccctgcacaataatccagaatggaatcccaactttaagaagaattatcgccttcacatgc 3203  
QY 3597 atcaatgtccacaataatggaatcacactactctgttagtaatgtgtcagagatccaagaatt 3656  
Db 3204 atcaatgtccacaataatggaatcacactactctgttagtaatgtgtcagagatccaagaatt 3263  
QY 3657 cgaatgtatctgtctcagaatggtgcagcaatgtaaaacatccatctcatcattctcagtga 3716  
Db 3264 cgaatgtatctgtctcagaatggtgcagcaatgtaaaacatccatctcatcattctcagtga 3323  
QY 3717 catgtgttcaactgtacgaaaaaagaagatataaatgtgcactgtacatctctatcca 3776  
Db 3324 catgtgttcaactgtacgaaaaaagaagatataaatgtgcactgtacatctctatcca 3383  
QY 3777 ggtgttttctgagacagtggaaatgttaccatccaaagctggaatttggcgggttggaatgc 3836  
Db 3384 ggtgttttctgagacagtggaaatgttaccatccaaagctggaatttggcgggttggaatgc 3443  
QY 3837 cttaattggcagacatctcatcactgtcgtggaatgagacacacttctctgtgttaacagcataag 3896  
Db 3444 cttaattggcagacatctcatcactgtcgtggaatgagacacacttctctgtgttaacagcataag 3503  
QY 3897 tgtcagaatccccctgggaatgtgctctctggaacacattagaattctcagaattacagcttca 3956  
Db 3504 tgtcagaatccccctgggaatgtgctctctggaacacattagaattctcagaattacagcttca 3563  
QY 3957 ggaacaatggaacagtgtggcccccaagaagtgcagacacttcaatccggatccaatcat 4016  
Db 3564 ggaacaatggaacagtgtggcccccaagaagtgcagacacttcaatccggatccaatcat 3623  
QY 4017 gacctgagacccaagaagcccttctctctggaatcaaggtggaatctgtgtggccaatgatt 4076  
Db 3624 gacctgagacccaagaagcccttctctctggaatcaaggtggaatctgtgtggccaatgatt 3683  
QY 4077 attcacggatccaagaccagggtgtcccgatcagaagtctctccagccttaacatctctag 4136  
Db 3684 attcacggatccaagaccagggtgtcccgatcagaagtctctccagccttaacatctctag 3743  
QY 4137 tttaatactatgtatagttctgtgaggaagaagtgcagagatctcgaggaattcccaact 4196  
Db 3744 tttaatactatgtatagttctgtgaggaagaagtgcagagatctcgaggaattcccaact 3803  
QY 4197 ggaaccttaatgtctctctctgtgcaatgtgatacatctggaataaacaacataatttt 4256  
Db 3804 ggaaccttaatgtctctctctgtgcaatgtgatacatctggaataaacaacataatttt 3863  
QY 4257 aaacctccaattatgtctgatacatccggttgcacccaactcatatagcatctgcagc 4316  
Db 3864 aaacctccaattatgtctgatacatccggttgcacccaactcatatagcatctgcagc 3923  
QY 4317 acctctgcagatgagtgagtgaggtgtgatttaatatagttgcacatgccaattgggaatg 4376  
Db 3924 acctctgcagatgagtgagtgaggtgtgatttaatatagttgcacatgccaattgggaatg 3983

QY 4377 gagagtaagaacatatacatgatagcacaagattactgtctcatcctactcttacaatatgttt 4436  
Db 3984 gagagtaagaacatatacatgatagcacaagattactgtctcatcctactcttacaatatgttt 4043  
QY 4437 gccacctgtctctctccaaaagctcagacttcaacctccaaggagagtaatgctctgaga 4496  
Db 4044 gccacctgtctctctccaaaagctcagacttcaacctccaaggagagtaatgctctgaga 4103  
QY 4497 cctcagggtgaataatccaaaagatggtgtcgaagtgaactccagaagaacatgaaagtgc 4556  
Db 4104 cctcagggtgaataatccaaaagatggtgtcgaagtgaactccagaagaacatgaaagtgc 4163  
QY 4557 acaggagtaactactcaaggagtaaaatctctgtcttaccagatgatagtgaaggagctc 4616  
Db 4164 acaggagtaactactcaaggagtaaaatctctgtcttaccagatgatagtgaaggagctc 4223  
QY 4617 ctcatctccagcagatccaagatggtgccatcagtgagactctcttttttcagaatgcaaatga 4676  
Db 4224 ctcatctccagcagatccaagatggtgccatcagtgagactctcttttttcagaatgcaaatga 4283  
QY 4677 aaggttttccagggaataatccaagactccttcaacactgtgtgtaactctcagaaccacg 4736  
Db 4284 aaggttttccagggaataatccaagactccttcaacactgtgtgtaactctcagaaccacg 4343  
QY 4737 ttactgactcgtactcctcgaattcaacccccagaatgtgggtgacacagattgacctgag 4796  
Db 4344 ttactgactcgtactcctcgaattcaacccccagaatgtgggtgacacagattgacctgag 4403  
QY 4797 atggaagttctgtggtcgtcgagagcagaagacctctcaactgactcgagctgc 4848  
Db 4404 atggaagttctgtggtcgtcgagagcagaagacctctcaactgaggtggtgcacatgc 4455

RESULT 8  
AAV19581  
ID AAV19581 standard; cDNA: 4832 BP.  
XX  
XX AAV19581;  
AC  
XX  
XX 06-AUG-1998 (first entry)  
DT  
XX  
XX Human factor VIII beta-domain deleted SQN deletion cDNA sequence.  
DE  
XX  
XX Replication defective; recombinant retrovirus; RRV; therapeutic protein;  
KW haemophilia; thrombosis; hypercoagulable disorder; liver disease; human;  
KW hepatitis; thalassemia; phenylketonuria; Lesch-Nyhan syndrome; diabetes;  
KW cystic fibrosis; Duchenne's Muscular Dystrophy; hypercholesterolemia;  
KW hypopituitarism; adenine deaminase deficiency; HIV infection; anaemia;  
KW Gucher's syndrome; high blood pressure; Alzheimer's disease, autoimmune;  
KW inflammatory disease; factor VIII; ss.  
XX  
XX Homo sapiens.  
OS  
XX  
XX W09800541-A2.  
FN  
XX  
XX 08-JAN-1998.  
PD  
XX  
XX 02-JUL-1997; 97WO-US11784.  
PE  
XX  
XX 04-JUN-1997; 97US-0869309.  
PR  
XX  
XX 03-JUL-1996; 96US-0645601.  
PR  
XX  
XX 13-AUG-1996; 96US-0696381.  
PA  
XX  
XX (CHIR ) CHIRON CORP.  
XX  
XX Allen JR, Barber JR, Boder M, Chang SM, Chang K;  
PI De LA VEGA D, Depoloni, Greengard J, Hsu DC, Ibanez CE;  
PI Jolly DJ, Lee R, Mittelstaedt DM, Prusak CE, Respass JG;  
XX  
XX WPI: 1998-086966/08.  
DR  
XX  
XX P-PSDB; AAW46246.  
XX







|||||  
Db 1691 aactaaatcagatcctcgttgccttgaccgcatctactctagtttcgttaatatggaag 1750  
QY 2099 agatctagcttcaggaactcatcttggtccctctctactctctgtaacaaagaatctgtatga 2158  
Db 1751 agatctagcttcaggaactcatcttggtccctctctactctctgtaacaaagaatctgtatga 1810  
QY 2159 aagaagaaaccaaataatgtctagaaagaagatgtcatacctgtttctctgtatactgtatga 2218  
Db 1811 aagaagaaaccaaataatgtctagaaagaagatgtcatacctgtttctctgtatactgtatga 1870  
QY 2219 gaaccgaagctgtgtaacctcacaagagaataatacaagctttctcccaatccagctggag 2278  
Db 1871 gaaccgaagctgtgtaacctcacaagagaataatacaagctttctcccaatccagctggag 1930  
QY 2279 gcaactggaagatccagaagttccaaagctccaaactcatgcaagcatcaaatggtatgt 2338  
Db 1931 gcaactggaagatccagaagttccaaagctccaaactcatgcaagcatcaaatggtatgt 1990  
QY 2339 ttctgatagtttgcagttgtctagtttgttgcatagagtgagatactgtatactctaa 2398  
Db 1991 ttctgatagtttgcagttgtctagtttgttgcatagagtgagatactgtatactctaa 2050  
QY 2399 catctggagcaagaactgtaactctctctctctctctctctctctctgatactcctcaacaa 2458  
Db 2051 catctggagcaagaactgtaactctctctctctctctctctctctgatactcctcaacaa 2110  
QY 2459 aatggtcatgtaagaacaactcacaactctccatctccaaagagaactgtctatgtc 2518  
Db 2111 aatggtcatgtaagaacaactcacaactctccatctccaaagagaactgtctatgtc 2170  
QY 2519 gatggaaaaaccaaagctctatgagatctggggtgcacaaactcaagaactctcgaaacaga 2578  
Db 2171 gatggaaaaaccaaagctctatgagatctggggtgcacaaactcaagaactctcgaaacaga 2230  
QY 2579 catgaccgctctactgnaagtttctagttgtgacaagaacactgtgtattatcaagaga 2638  
Db 2231 catgaccgctctactgnaagtttctagttgtgacaagaacactgtgtattatcaagaga 2290  
QY 2639 cagttatgaagaatacttcagatactgtctgataaacaatgcatctgaaacaaag 2698  
Db 2291 cagttatgaagaatacttcagatactgtctgataaacaatgcatctgaaacaaag 2350  
QY 2699 ctctcccaagaatccaagaacacctagcaactagcgaatgaattaaagccccacc 2758  
Db 2351 ctctcccaaga-----ccccacc 2368  
QY 2759 agtcttgaacgcatacaacgggaataactcgtactactcttcagttcaagatcaagaga 2818  
Db 2369 agtcttgaacgcatacaacgggaataactcgtactactcttcagttcaagatcaagaga 2428  
QY 2819 aatgtaataatgataacatacatagttgaaatgaaagaagaatttgacattatga 2878  
Db 2429 aatgtaataatgataacatacatagttgaaatgaaagaagaatttgacattatga 2488  
QY 2879 tgaagatgaaaaatcaagccccgcagcttccaagaagaacacatacttatgtc 2938  
Db 2489 tgaagatgaaaaatcaagccccgcagcttccaagaagaacacatacttatgtc 2548  
QY 2939 tgaagatgaaagctctggggttataggatagtagctcccaactgttttaagaagaag 2998  
Db 2549 tgaagatgaaagctctggggttataggatagtagctcccaactgttttaagaagaag 2608  
QY 2999 ggtcacaagtgcagttcctcctcagttcaagaagaattgtttccaggaatttaccatg 3058  
Db 2609 ggtcacaagtgcagttcctcctcagttcaagaagaattgtttccaggaatttaccatg 2668  
QY 3059 ctctcttactcagccctataccgttggaagaactaaatgaacatttgggaactcttgggccc 3118  
Db 2669 ctctcttactcagccctataccgttggaagaactaaatgaacatttgggaactcttgggccc 2728  
QY 3119 atataataagcagaagtgtgaagataataataatgtaacttcaagaatccagctctcg 3178  
|||||

Db 2729 atataaagcagaagttgaagataataataatgtaacttcaagaatccagctctcg 2788  
QY 3179 tcccatctctctacttcttaagcttattcttatbaagaaatcagaagcagaagcaga 3238  
Db 2789 tcccatctctctacttcttaagcttattcttatbaagaaatcagaagcagaagcaga 2848  
QY 3239 accatgaaaaaacttgttcaagcctaataagaaacaaactacttltggaagtgcaca 3298  
Db 2849 accatgaaaaaacttgttcaagcctaataagaaacaaactacttltggaagtgcaca 2908  
QY 3299 tcatatggaacccactaaagatgaattgtgactgcaaaacccgggtctattctctgagt 3358  
Db 2909 tcatatggaacccactaaagatgaattgtgactgcaaaacccgggtctattctctgagt 2968  
QY 3359 tgaactggaaaaaagatgtgcaactcagcctgattggaacccctctgctgcacactaa 3418  
Db 2969 tgaactggaaaaaagatgtgcaactcagcctgattggaacccctctgctgcacactaa 3028  
QY 3419 cacaatgaaacccctgctcatgaggagaacaagtgaagtaacagaattgtctgttttaac 3478  
Db 3029 cacaatgaaacccctgctcatgaggagaacaagtgaagtaacagaattgtctgttttaac 3088  
QY 3479 catctttgataagaccaaagctgtgtacttcaactgaaataatggaagaacatgcaggc 3538  
Db 3089 catctttgataagaccaaagctgtgtacttcaactgaaataatggaagaacatgcaggc 3148  
QY 3539 tccctgcaataatccagatggaagatcccaactttaaagaatatactgctccatgcaat 3598  
Db 3149 tccctgcaataatccagatggaagatcccaactttaaagaatatactgctccatgcaat 3208  
QY 3599 caatggtacataatgataatgatacaactcgtgcttagtaatggtcgaagatcaagattcg 3658  
Db 3209 caatggtacataatgataatgatacaactcgtgcttagtaatggtcgaagatcaagattcg 3268  
QY 3659 atggtatctgctcagacatggggcagaatgaaacaatccatctcttcaattcaatgaga 3718  
Db 3269 atggtatctgctcagacatggggcagaatgaaacaatccatctcttcaattcaatgaga 3328  
QY 3719 tgtgtcactgtaacgaaaaaaagaggagtalaaatgtaacatgtaacatctctacag 3778  
Db 3329 tgtgtcactgtaacgaaaaaaagaggagtalaaatgtaacatgtaacatctctacag 3388  
QY 3779 tgtttttgaagacagtgtgaatgttaccatccaagctvgaaattvggggtgtgaatgct 3838  
Db 3389 tgtttttgaagacagtgtgaatgttaccatccaagctvgaaattvggggtgtgaatgct 3448  
QY 3839 tatltggagaactctacatgctggggtagaagacacttttctgtgtgaacgaataagt 3898  
Db 3449 tatltggagaactctacatgctggggtagaagacacttttctgtgtgaacgaataagt 3508  
QY 3899 tcaagactccctctgggaatggtctctggaacacattagagatttcaagattcaagctcag 3958  
Db 3509 tcaagactccctctgggaatggtctctggaacacattagagatttcaagattcaagctcag 3568  
QY 3959 acaataatggaacagtgtgcccccaagctvgccaagacttcaattatccggatcaatcatgc 4018  
Db 3569 acaataatggaacagtgtgcccccaagctvgccaagacttcaattatccggatcaatcatgc 3628  
QY 4019 ctggagacccaagagaccccttcttggatacaaggtgatactgttggcaacaaatgattat 4078  
Db 3629 ctggagacccaagagaccccttcttggatacaaggtgatactgttggcaacaaatgattat 3668  
QY 4079 tcaagcgaatcaagaccaggtgtcccgtaagaatttccacagcctctcaactctcagtt 4138  
Db 3689 tcaagcgaatcaagaccaggtgtcccgtaagaatttccacagcctctcaactctcagtt 3748  
QY 4139 tatcatatgataagcttgaatggaagaagtggaagacttatactgaggaatctccactg 4198  
Db 3749 tatcatatgataagcttgaatggaagaagtggaagacttatactgaggaatctccactg 3808  
QY 4199 aaccttaatgctctctcttggcaatggtgatactctggaataaacaataatttttaa 4258  
Db 3809 aaccttaatgctctctcttggcaatggtgatactctggaataaacaataatttttaa 3868  
|||||



OY	4259	ccccccaattatgtcgtacatcacccgtttgcacccaactcttaataagatctcgagac	4318
Db	3869	cccccaattattgtcgtacatcacccgtttgcacccaactcttaataagatctcgagac	3928
OY	4319	tcttcgcatgtgagtttgatggcctgctgatttaaatagttgcagcatgccaatgaggaaatga	4378
Db	3929	tcttcgcatgtgagtttgatggcctgctgatttaaatagttgcagcatgccaatgaggaaatga	3988
OY	4379	gagtaagaagcaatatcagatgcacagagtttaactcttcacctaactcttaccaatatggttgc	4438
Db	3969	gagtaagaagcaatatcagatgcacagagtttaactcttcacctaactcttaccaatatggttgc	4048
OY	4439	cacctggtctctctcctcaaaaagctcgcgaacttcaactccaaggaggagtaaatgcttgagacc	4498
Db	4049	cacctggtctctctcctcctcaaaaagctcgcgaacttcaactccaaggaggagtaaatgcttgagacc	4108
OY	4499	tcaaagtgaataatccaaaagagatggctgcgaatgagacttccagaagacaatgaaagtac	4558
Db	4109	tcaaagtgaataatccaaaagagatggctgcgaatgagacttccagaagacaatgaaagtac	4168
OY	4559	aggaggtactactcgaaggagtaaaatctctgtcttccacgaatgtaatgtgaaggagttcct	4618
Db	4169	aggaggtactactcgaaggagtaaaatctctgtcttccacgaatgtaatgtgaaggagttcct	4228
OY	4619	catctccagcagatcgaagatgagccaatcagatgtagactctcttttttcagaatgycaaagttaa	4678
Db	4229	catctccagcagatcgaagatgagccaatcagatgtagactctctcttttttcagaatgycaaagttaa	4288
OY	4679	ggtttttcagggaataatcagaagactccttcaactctgtgtgaactctcagaaccacagctt	4738
Db	4289	ggtttttcagggaataatcagaagactccttcaactctgtgtgaactctcagaaccacagctt	4348
OY	4739	actgaactcgtctactcttcgaattcacccccaagattcgggtgcacagatctgcctgagagt	4798
Db	4349	actgaactcgtctactcttcgaattcacccccaagattcgggtgcacagatctgcctgagagt	4408
OY	4799	ggagagttctgtggtctgcgaggcacaggaactctactgaactgcagcgagttc	4848
Db	4409	ggagagttctgtggtctgcgaggcacaggaactctactgaaggttgcacatgc	4458
RESULT 9			
ID	AAV15338	standard; DNA; 4832 BP.	
XX	AAV15338;		
XX	20-JUL-1998	(first entry)	
DE	Human Factor VIII SQN deletion mutant DNA.		
KW	Factor VIII; blood clotting; haemophilia A; gene therapy;		
OS	Homo sapiens.		
OS	Synthetic.		
XX	Key	Location/Qualifiers	
FT	CDS	72..4445	
XX		/*tag= a	
PN	M09800542-A2.		
XX			
PD	08-JAN-1998.		
XX			
PE	02-JUL-1997;	97MO-US11785.	
XX			
PR	04-JUN-1997;	97US-0869309.	
PR	03-JUL-1996;	96US-0643601.	
PR	13-AUG-1996;	96US-0696381.	
PA	(CHIR ) CHIRON CORP.		

[illegible]



QY	959	tgvgagccctggtataaagacttgaatttcagagccctacattggagccctactagatgataagata	1018
Db	611	tgvgagccctggtataaagacttgaatttcagagccctacattggagccctactagatgataagata	670
QY	1019	aggagatctgvgccaagaagaaagacacagaccctgacacaagtta tactactcttgccgt 	1078
Db	671	aggagatctgvgccaagaagaaagacacagaccctgacacaatttactactcttgccgt	730
QY	1079	atttgaatgaaggagaaaaagtctgacatcagaaaacaagaactccttgaatgacagataagga	1138
Db	731	atttgaatgaaggagaaaaagtctgacatcagaaaacaagaactccttgaatgacagataagga	790
QY	1139	tgtgtgacatctgcctcgvgccctcgvgccttaaaatgacacacgtcaatgtgttatgtaaacaggtc	1198
Db	791	tgtgtgacatctgcctcgvgccctcgvgccttaaaatgacacacgtcaatgtgttatgtaaacaggtc	850
QY	1199	tctgcacagagctgcatatgtagatgccaagaagaaatcaagcttatgttgagatgtagatggag	1258
Db	851	tctgcacagagctgcatatgtagatgccaagaagaaatcaagcttatgttgagatgtagatggag	910
QY	1259	caacacatctcgaagatgacacccaatatctccctcgaaagtcacacacatctcttctgtagaagaa	1318
Db	911	caacacatctcgaagatgacacccaatatctccctcgaaagtcacacacatctcttctgtagaagaa	970
QY	1319	tgcgcagcgctctctggaatatctgcgccaataacttccctactgctccaacaacatcttgat	1378
Db	971	tgcgcagcgctctctctggaatatctgcgccaataacttccctactgctccaacaacatcttgat	1030
QY	1379	ggagcccttggaacgaatttctactggtttgtcaatctctccacccaagaatgtagcatgga	1438
Db	1031	ggagcccttggaacgaatttctactggtttgtcaatctctccacccaagaatgtagcatgga	1090
QY	1439	agccttatgtcaaaagttagacacgcgtgtccagaggaaccccaactaagaaatgaaataaataga	1498
Db	1091	agccttatgtcaaaagttagacacgcgtgtccagaggaaccccaactaagaaatgaaataaataga	1150
QY	1499	agaaagcggaaagactatgatagtacttactgacttctgaatggaatggtgtgtgtcaggtttga	1558
Db	1151	agaaagcggaaagactatgatagtacttactgacttctgaatggaatggtgtgtgtcaggtttga	1210
QY	1559	tgaatgaacactctccctcccttcttcacaaattcgttgtagttgccaagaagacatccctaaac	1618
Db	1211	tgaatgaacactctccctcccttcttcacaaattcgttgtagttgccaagaagacatccctaaac	1270
QY	1619	ttgggtgacatlaacatctgtctgtcgaagagggagacgtggagacatgactcccttagcctctgc	1678
Db	1271	ttgggtgacatlaacatctgtctgtcgaagagggagacgtggagacatgactcccttagcctctgc	1330
QY	1679	ccccgatgacagaagttataaaagccaatatctgaaacaatgccccctcagcggatgtgtag	1738
Db	1331	ccccgatgacagaagttataaaagccaatatctgaaacaatgccccctcagcggatgtgtag	1390
QY	1739	gaagtacaaaagaatctcogatttatatggaatacagaatgaacactttagactctgtgaagc	1798
Db	1391	gaagtacaaaagaatctcogatttatatggaatacagaatgaacactttagactctgtgaagc	1450
QY	1799	tattcagacatgaatcaggaatctctggagaccttbaacttaatggggaaagtgtgagacacact	1858
Db	1451	tattcagacatgaatcaggaatctctggagaccttbaacttaatggggaaagtgtgagacacact	1510
QY	1859	gttgatataatttaagaatatcaagaacagacagacatataaactctacccctcagcgaatatcac	1918
Db	1511	gttgatataatttaagaatatcaagaacagacacatataaactctacccctcagcgaatatcac	1570
QY	1919	tgaatctcgatcctcttgttatctcaagagagattacccaagaagtgttaaacatttgaagaattc	1978
Db	1571	tgaatctcgatcctcttgttatctcaagagagattacccaagaagtgttaaacatttgaagaattc	1630
QY	1979	tccaatctctcgagggagaaatattcaaaatbaaaatgagacgttgactgtagaagaattggcc	2038
Db	1631	tccaatctctcgagggagaaatattcaaaatbaaaatgagacgttgactgtagaagaattggcc	1690
QY	2039	aactaaatcagatccctctcggtgcctgaccgcgtatctactctagttctgtttaataatgagag	2098

Dd	1651	aactaaacagatccctcggcgcccgagccgcgtatctctcaagttcttgtaataatgagag	1750
Qy	2099	agatctaacctcgaagactcaatttgccctctccctcaatctctcaagaagatctgtagaca	2158
Dd	1751	agatctagcttcgagactcaattgctccctccctccatctgtctaaagaatctgtatgaca	1810
Qy	2159	aagaggaacacagataatgtctcagacaagaggaatgtcaacctgttctctgtattgtatga	2218
Dd	1811	aagaggaacacagataatgtctcagacaagaggaatgtcaacctgttctctgtattgtatga	1870
Qy	2219	gaaccggaagctggtacctcaacaggaatatacaagccttctcccaatccagctgaggt	2278
Dd	1871	gaaccggaagctggtacctcaacaggaatatacaagccttctcccaatccagctgaggt	1930
Qy	2279	gcaagcttaagatcccaagatgtcccaagcctccaacatcatgtcaagaatcaatgtcgtagt	2338
Dd	1931	gcaagcttaagatcccaagatgtcccaagcctccaacatcatgtcaagaatcaatgtcgtagt	1990
Qy	2339	ttttgatagttttgacgttgtcaagttgtttgtgca tgaagtgagcaatggtatcatcttaag	2398
Dd	1991	ttttgatagttttgacgttgtcaagttgtttgtgca tgaagtgagcaatggtatcatcttaag	2050
Qy	2399	catgttgagcaagaagctgaactctctctgtctctctctctctctgtatataacttcaaacaa	2458
Dd	2051	catgttgagcaagaagctgaactctctctgtctctctctctctctgtatataacttcaaacaa	2110
Qy	2459	aatgtctcattgaagaacacactccaacccatctcccatctctcaagagaagaactgtctatgtc	2518
Dd	2111	aatgtctcattgaagaacacactccaacccatctcccatctctcaagagaagaactgtctatgtc	2170
Qy	2519	gattgaaaaacccaaggtctatgtatgtattctgggtgtgccaaactagaactctggagaacagag	2578
Dd	2171	gattgaaaaacccaaggtctatgtatgtattctgggtgtgccaaactagaactctggagaacagag	2230
Qy	2579	catgtaccgcttactgtgaaggttctctagttgtgtgcaagaacactggtgtattatcagaga	2638
Dd	2231	catgtaccgcttactgtgaaggttctctagttgtgtgcaagaacactggtgtattatcagaga	2290
Qy	2639	caattatgaataatattcttcagcaactgtgcagatgaanaacatgtccatgtgaacccaagag	2698
Dd	2291	caattatgaataatattcttcagcaactgtgcagatgaanaacatgtccatgtgaacccaagag	2350
Qy	2699	cttctccccaagaattcaaagacacccctagcacctagggcaaaaagcaatttaatgtccaccac	2758
Dd	2351	cttctccccaagaattcaaagacacacccctagcacctagggcaaaaagcaatttaatgtccaccac	2368
Qy	2759	agctcttgaaacgcacccaagaagggaataaactgtactactcttcagtcagatcacaagaga	2818
Dd	2369	agctcttgaaacgcacccaagaagggaataaactgtactactcttcagtcagatcacaagaga	2428
Qy	2819	aattggaactatgatatccatcatcaatcaagtgttgaaatgaagaaggaagattttgacatttga	2878
Dd	2429	aattggaactatgatatccatcatcaatcaatcaagtgttgaaatgaagaaggaagattttgacatttga	2488
Qy	2879	tgaaagtatgaataatcagaagcccccgcagccttccaaaagaagaacagacatattattatgc	2938
Dd	2489	tgaaagtatgaataatcagaagcccccgcagccttccaaaagaagaacagacatattattatgc	2548
Qy	2939	tgcaagtgtgaagaagctctgtggattatgtggaatgagtgatcccccacaatgtcttaagaacag	2998
Dd	2549	tgcaagtgtgaagaagctctgtggattatgtggaatgagtgatcccccacaatgtcttaagaacag	2608
Qy	2999	ggctccagaagtgtgcaggtgtccctcaagttccaagaaagtgttttccaggaattatcagtagg	3058
Dd	2609	ggctccagaagtgtgcaggtgtccctcaagttccaagaaagtgttttccaggaattatcagtagg	2668
Qy	3059	ctctcttaactcaagccctataacgtgtgagaacccaattgaaactttgggaactccctggggcc	3118
Dd	2669	ctctcttaactcaagccctataacgtgtgagaacccaattgaaactttgggaactccctggggcc	2728
Qy	3119	atatataagagccaagaatttgaagaataatcatcgtgtaacttccagaatccagcctctcgg	3178



Db 2729 atataagagcagaagttgagaataatcatgtgtaacttcagaaatcagccctctcg 2788  
 QY 3179 tccctattccttattcttaagccttatttcttatgaggaagatcagaggcagaagcaga 3238  
 Db 2789 tccctattccttattcttaagccttatttcttatgaggaagatcagaggcagaagcaga 2848  
 QY 3239 acccagaaaactttgcaagcctaatagaacaaacttactttggaagatgcaaca 3298  
 Db 2849 acccagaaaactttgcaagcctaatagaacaaacttactttggaagatgcaaca 2908  
 QY 3299 tcataatgaccccaataagaatgagttgactgcaagcctgggctatttctctgagt 3358  
 Db 2909 tcataatgaccccaataagaatgagttgactgcaagcctgggctatttctctgagt 2968  
 QY 3359 tgacctggaaaagaatgtgcactcagcgctgagtggagcccttctgtctgcaactaa 3418  
 Db 2969 tgacctggaaaagaatgtgcactcagcgctgagtggagcccttctgtctgcaactaa 3028  
 QY 3419 cacactgaacctgtcatatggagacagaagtgacagtaacgaattgctctgttttcac 3478  
 Db 3029 cacactgaacctgtcatatggagacagaagtgacagtaacgaattgctctgttttcac 3088  
 QY 3479 catctttgatagagaccaaagctgtgtaacttcaactgaaaataatggaaaagaaactgcaaggc 3538  
 Db 3089 catctttgatagagaccaaagctgtgtaacttcaactgaaaataatggaaaagaaactgcaaggc 3148  
 QY 3539 tccctgcaatatccagatgtgagaatcccaactttaaagagatatatgcttccatgcaat 3598  
 Db 3149 tccctgcaatatccagatgtgagaatcccaactttaaagagatatatgcttccatgcaat 3208  
 QY 3599 caatgctacataatgtgatacacactacctgtgcttagtaatgctcagagatcaaaagatctcg 3658  
 Db 3209 caatgctacataatgtgatacacactacctgtgcttagtaatgctcagagatcaaaagatctcg 3268  
 QY 3659 atggatctgtctcagcattggtgagcaaatgaaaacatccatctatctcaattcagttgaca 3718  
 Db 3269 atggatctgtctcagcattggtgagcaaatgaaaacatccatctatctcaattcagttgaca 3328  
 QY 3719 tgtgttactgtttagaaaagaagaggttaaaatgagcctgtgacaactctcttccag 3778  
 Db 3229 tgtgttactgtttagaaaagaagaggttaaaatgagcctgtgacaactctcttccag 3388  
 QY 3779 tgtttttgagcagtggaatgtttacatcccaagctggaatttggcggtgtgaaatgct 3838  
 Db 3389 tgtttttgagcagtggaatgtttacatcccaagctggaatttggcggtgtgaaatgct 3448  
 QY 3839 taattggcgagcatctacatgctgtggatgagcaacttcttctgtgtacagcaataagtg 3898  
 Db 3449 taattggcgagcatctacatgctgtggatgagcaacttcttctgtgtgtacagcaataagtg 3508  
 QY 3899 tcagactcccttggaatgtgcttctgacacatagaagatttaagaatttcaagcttcag 3958  
 Db 3509 tcagactcccttggaatgtgcttctgacacatagaagatttaagaatttcaagcttcag 3568  
 QY 3959 acaatatgagcagtggtggtcccaaaagctggtgcaagatcatatctcgatcaatcaatg 4018  
 Db 3569 acaatatgagcagtggtggtcccaaaagctggtgcaagatcatatctcgatcaatcaatg 3628  
 QY 4019 ctggagcacaagaagccttcttctgtgataaagtgagatctgttggcaccaatgattat 4078  
 Db 3629 ctggagcacaagaagccttcttctgtgataaagtgagatctgttggcaccaatgattat 3688  
 QY 4079 tcagagcacaagaagcgtggtggttcaggaagtctcagagctctacactcttcaggt 4138  
 Db 3689 tcagagcacaagaagcgtggtggttcaggaagtctcagagctctacactcttcaggt 3748  
 QY 4139 tatcatatgtatagttctatgaggaagaagtggcagactatctgagaagaattccactgg 4198  
 Db 3749 tatcatatgtatagttctatgaggaagaagtggcagactatctgagaagaattccactgg 3808  
 QY 4199 aaccttaatgtcttcttctgtgcaatgtgagttcaactgtggataaacaacataatttaa 4258  
 Db 3809 aaccttaatgtcttcttctgtgcaatgtgagttcaactgtggataaacaacataatttaa 3868

QY 4259 cctcccaattatgtctgatatacatcgtttgaccccaactcatratatgacttgcagcac 4318  
 Db 3869 cctcccaattatgtctgatatacatcgtttgaccccaactcatratatgacttgcagcac 3928  
 QY 4319 tcttcgcatgtagtctgtaggtgtgatttaaatagttgtgcagatgcatatggaaatgga 4378  
 Db 3929 tcttcgcatgtagtctgtaggtgtgatttaaatagttgtgcagatgcatatggaaatgga 3988  
 QY 4379 gagtaagcaatatcagaatgacagatbaactgtcttcaacttcaacttcaactatgtttgc 4438  
 Db 3989 gagtaagcaatatcagaatgacagatbaactgtcttcaacttcaacttcaactatgtttgc 4048  
 QY 4439 cactgtgtctcttcccaaaagctgagatccacttccaaaggagagatgactgtgagacc 4498  
 Db 4049 cactgtgtctcttcccaaaagctgagatccacttccaaaggagagatgactgtgagacc 4108  
 QY 4499 tcaggtgaataatccaaaaggtgtgcaagtgtgacttccagaagaacatgaaatgcaac 4558  
 Db 4109 tcaggtgaataatccaaaaggtgtgcaagtgtgacttccagaagaacatgaaatgcaac 4168  
 QY 4559 aggaagtaactactcagggaggttaaaatctctgtcttaccagcatgtatgtgaagagttcct 4618  
 Db 4169 aggaagtaactactcagggaggttaaaatctctgtcttaccagcatgtatgtgaagagttcct 4228  
 QY 4619 catctccagcagtgaaatgtgcacatcagtgactctcttcttgaagatggcaagaatgaa 4678  
 Db 4229 catctccagcagtgaaatgtgcacatcagtgactctcttcttgaagatggcaagaatgaa 4288  
 QY 4679 gtttttcaggaagaatcaagaactccttcaacacctgtgtgtaactctcttgaagccacggt 4738  
 Db 4289 gtttttcaggaagaatcaagaactccttcaacacctgtgtgtaactctcttgaagccacggt 4348  
 QY 4739 actgactgtactacttgcgaattcaacccccagagttggtgacacagattgctcctgagat 4798  
 Db 4349 actgactgtactacttgcgaattcaacccccagagttggtgacacagattgctcctgagat 4408  
 QY 4799 ggaagttctgtggtgtgcagagcagagacacttactgagaggtgtgacactgc 4848  
 Db 4409 ggaagttctgtggtgtgcagagcagagacacttactgagaggtgtgacactgc 4458  
 RESULT 10  
 AAA49232  
 ID AAA49232 standard; DNA: 12445 BP.  
 XX  
 AC AAA49232;  
 XX  
 DT 26-SEP-2000 (first entry)  
 XX  
 DE Vector HSQReNeo for transforming endothelial cells.  
 XX  
 KW Hemostatic; antianemic; antidiabetic; nootropic; neuroprotective; vector;  
 KW osteopathic; antislaking; immunostimulant; gene therapy; collagen;  
 KW endothelial cell; peripheral blood; buffy coat cell; VEGF; PCR primer;  
 KW vascular endothelial growth factor; bovine brain extract; haemophilia;  
 KW Factor VIII; human; transgene; adenosine deaminase deficiency; ss;  
 KW sickle cell anaemia; thalassemia; diabetes; alpha-antitrypsin deficiency;  
 KW Alzheimer's disease; brain disease; heart disease; immune system defect;  
 KW bone fracture; osteoporosis.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 XX  
 PN WO200032750-A1.  
 XX  
 PD 08-JUN-2000.  
 XX  
 PF 24-NOV-1999; 99MO-US28033.  
 XX  
 PR 24-NOV-1998; 98US-0109687.  
 XX  
 PA (MINN ) UNIV MINNESOTA.



PA (UYEM-) UNIV EMORY.  
PA (HEBB/) HEBBEL R P.  
PA (LINY/) LIN Y.  
PA (LOLL/) LOLLAR J S.  
XX  
PI Hebbel RP, Lin Y, Lollar JS:  
XX  
DR WPI; 2000-412303/35.  
XX  
PT Expanding population of endothelial cells useful to biocompatibilize  
PT implantable medical devices comprises contacting buffy coat cells with  
PT collagen I coated surtice in culture medium comprising vascular  
PT endothelial growth factor -  
XX  
PS Claim 19; Fig 4; 53pp; English.  
XX  
CC The invention relates to a method for expanding the population of  
CC endothelial cells (EC) obtained from peripheral blood, by culturing,  
CC in contact with a collagen I coated surface, buffy coat cells obtained  
CC from peripheral mammalian blood in the presence of a culture medium  
CC containing vascular endothelial growth factor (VEGF) and free of bovine  
CC brain extract. EC are useful for treating hemophilia by introducing it  
CC into the blood stream of a mammal, so that an effective amount of  
CC Factor VIII protein is secreted in the blood stream of the mammal.  
CC This sequence represents the vector HSPkneo where the human factor  
CC VIII gene into which an enhanced green fluorescent protein coding  
CC sequence (HSG) has been inserted, is subcloned. Transgenic EC transduced  
CC in vitro are useful for improving prosthetic implants. EC is also useful  
CC for diagnosing clotting disorders where indication or disease is  
CC associated with a reduction in the activity of an enzyme. EC is also  
CC useful in gene therapy for treating the variety of diseases including  
CC adenosine deaminase deficiency, sickle cell anemia, thalassemia,  
CC hemophilia, diabetes, alpha-antitrypsin deficiency, brain disorders  
CC such as Alzheimer's disease, heart diseases, defects in immune system,  
CC for repairing bone fractures and to treat or prevent osteoporosis.  
XX  
SQ Sequence 12445 BP; 3255 A; 2989 C; 2951 G; 3250 T; 0 other;

Query Match 54.4%; Score 4319; DB 21; Length 12445;  
Best Local Similarity 98.9%; Pred. No. 0;  
Matches 4374; Conservative 0; Mismatches 5; Indels 42; Gaps 1;

QY 415 ccacatgataatagagctctccacgtctctctctgctctgctctgcatctctta 474  
DB |||||||  
DB 763 ccacatgataatagagctctccacgtctctctctgctctgctctgcatctctta 822  
QY 475 gtgcacacagaagatctactcgtgtgagtggaactgcatggacatatacgcaagt 534  
DB |||||||  
DB 823 gtgcacacagaagatctactcgtgtgagtggaactgcatggacatatacgcaagt 882  
QY 535 atctcgtgtaagctcgtgtgagtggaactgcatggacatatacgcaagt 594  
DB |||||||  
DB 883 atctcgtgtaagctcgtgtgagtggaactgcatggacatatacgcaagt 942  
QY 595 tcaaacctcagtcgtgtacaagaagactcgtgtgtagaattcaagtgtaacatttca 654  
DB |||||||  
DB 943 tcaaacctcagtcgtgtacaagaagactcgtgtgtagaattcaagtgtaacatttca 1002  
QY 655 acatcgttaaacgaagccctcgtgagtggtctcgtagtcctcctccacccgggtgag 714  
DB |||||||  
DB 1003 acatcgttaaacgaagccctcgtgagtggtctcgtagtcctcctccacccgggtgag 1062  
QY 715 ttctatgatacgtgtgcatctacttaagaacatgctccacatctgtaacttcaatg 774  
DB |||||||  
DB 1063 ttctatgatacgtgtgcatctacttaagaacatgctccacatctgtaacttcaatg 1122  
QY 775 ctgtgtgtatctactactggaagctctcgtgagtggaactgtaataatgtaacagacagtc 834  
DB |||||||  
DB 1123 ctgtgtgtatctactactggaagctctcgtgagtggaactgtaataatgtaacagacagtc 1182  
QY 835 aaaggggaaagaagatgataaagctctcctcgtgtggaagcacaatcatatgctgtgcaag 894  
DB |||||||

DB 1183 aaaggggaaagaagatgataaagctctcctcgtgtggaagcacaatcatatgctgtgcaag 1242  
QY 895 tccctgaagaagatggtccaatgagctctgacccactgtgcttaaccatactctt 954  
DB |||||||  
DB 1243 tccctgaagaagatggtccaatgagctctgacccactgtgcttaaccatactctt 1302  
QY 955 ctcatgtgacctgtgtaaagaacttgaattcaggccctatgtagccctactagatgta 1014  
DB |||||||  
DB 1303 ctcatgtgacctgtgtaaagaacttgaattcaggccctatgtagccctactagatgta 1362  
QY 1015 gagaaggagctctggtccaaagaaagacacagacttgcacaattatctactctttg 1074  
DB |||||||  
DB 1363 gagaaggagctctggtccaaagaaagacacagacttgcacaattatctactctttg 1422  
QY 1075 ctgtatgtatgaaggaaagttgacatcgaagaacaaagacactctgtatgagata 1134  
DB |||||||  
DB 1423 ctgtatgtatgaaggaaagttgacatcgaagaacaaagacactctgtatgagata 1482  
QY 1135 ggaatgtgcatctgtctcggtcgtgcttaaatgacacagtcgaatgtatgtaaca 1194  
DB |||||||  
DB 1483 ggaatgtgcatctgtctcggtcgtgcttaaatgacacagtcgaatgtatgtaaca 1542  
QY 1195 ggtctctgcagtgctgtatgtgtagtcgaagaaatcagtcctatgtgtagttgaa 1254  
DB |||||||  
DB 1543 ggtctctgcagtgctgtatgtgtagtcgaagaaatcagtcctatgtgtagttgaa 1602  
QY 1255 tgggacacacctcctgaatgacatcaatatctctgaagtgacacacatctcttgagaa 1314  
DB |||||||  
DB 1603 tgggacacacctcctgaatgacatcaatatctctgaagtgacacacatctcttgagaa 1662  
QY 1315 accatcgcagagcgctccttggaatctgcgaataacttccctactgcttaaacactct 1374  
DB |||||||  
DB 1663 accatcgcagagcgctccttggaatctgcgaataacttccctactgcttaaacactct 1722  
QY 1375 tgatgacacttggaagttctactcgtgttgcatatctctccacaacatgatggca 1434  
DB |||||||  
DB 1723 tgatgacacttggaagttctactcgtgttgcatatctctccacaacatgatggca 1782  
QY 1435 tggaaagctatgttcaaaatagacaagctgtgccagagaaacccaactaagaaatgaaata 1494  
DB |||||||  
DB 1783 tggaaagctatgttcaaaatagacaagctgtgccagagaaacccaactaagaaatgaaata 1842  
QY 1495 atgaaagaagcggaaagactatgtagatctctactgattcggaaatgtagtgcaggt 1554  
DB |||||||  
DB 1843 atgaaagaagcggaaagactatgtagatctctactgattcggaaatgtagtgcaggt 1902  
QY 1555 ttgatgatgaacaactctcctctcttatccaattcgtctcagttgccaagaagatccta 1614  
DB |||||||  
DB 1903 ttgatgatgaacaactctcctctcttatccaattcgtctcagttgccaagaagatccta 1962  
QY 1615 aaacttgggtacatatactgtctgtgaagaggaagacttggacatagctctctagctc 1674  
DB |||||||  
DB 1963 aaacttgggtacatatactgtctgtgaagaggaagacttggacatagctctctagctc 2022  
QY 1675 tgcgcccgatgacagaagttataaagttcaatatcttgaacaaatgagccctcgaagtgatg 1734  
DB |||||||  
DB 2023 tgcgcccgatgacagaagttataaagttcaatatcttgaacaaatgagccctcgaagtgatg 2082  
QY 1735 gtgagaagtacaaaagaatccgatattgycatacaagatggaacacttgaagctcgtg 1794  
DB |||||||  
DB 2083 gtgagaagtacaaaagaatccgatattgycatacaagatggaacacttgaagctcgtg 2142  
QY 1795 aagcatcagcatgaatcagaagatctctggaacttacttattggtggaagtgtgagaca 1854  
DB |||||||  
DB 2143 aagcatcagcatgaatcagaagatctctggaacttacttattggtggaagtgtgagaca 2202  
QY 1855 caagtgtatataatcttaagaatcaagcaagcagacatatatactaccctcaagaa 1914  
DB |||||||  
DB 2203 caagtgtatataatcttaagaatcaagcaagcagacatatatactaccctcaagaa 2262  
QY 1915 tcaagtgtatcgtctcttattatcaagaagattacaaaagtggttaaaacttggagg 1974  
DB |||||||  
DB 2263 tcaagtgtatcgtctcttattatcaagaagattacaaaagtggttaaaacttggagg 2322



QY	1975	atttccaatcttcgcccaggagaatatctcaaatataaatgagcagctgactctgtagaagctg	2034
Db	2333	atttccaatcttcgcccaggagaatatctcaaatataaatgagcagctgactctgtagaagctg	2382
QY	2035	ggccaactaaatcagatccctccgctgctcgtaccgcgcattactacgttctgcttaatgag	2094
Db	2383	ggccaactaaatcagatccctccgctgctcgtaccgcgcattactacgttctgcttaatgag	2442
QY	2095	agagagatctagcttcaaggacatctgcccctccctcaatcgtctacaagaatctgtag	2154
Db	2443	agagagatctagcttcaaggacatctgcccctccctcaatcgtctacaagaatctgtag	2502
QY	2155	atcaaaagagaaaccaaataatgtcaagaaagaagatgtactccgtttctgtatttg	2214
Db	2503	atcaaaagagaaaccaaataatgtcaagaaagaagatgtactccgtttctgtatttg	2562
QY	2215	atggaacccgaagctctgttaactccacagagaataataaagccttccccaatccagctg	2274
Db	2563	atggaacccgaagctctgttaactccacagagaataataaagccttccccaatccagctg	2622
QY	2275	gagctgcagctctgagatccagagcttccaaagcctccaacaatctgcacagcatcaatgct	2334
Db	2623	gagctgcagctctgagatccagagcttccaaagcctccaacaatctgcacagcatcaatgct	2682
QY	2335	atgtttttgatagttttcagcttgcagctgtgtttgtgatbagctgcatctacgtacatc	2394
Db	2683	atgtttttgatagttttcagcttgcagctgtgtttgtgatbagctgcatctacgtacatc	2742
QY	2395	taagcatctggagcaacagacttccctctctgctctctctctctcgtatataccttcaac	2454
Db	2743	taagcatctggagcaacagacttccctctctgctctctctctctcgtatataccttcaac	2802
QY	2455	acaaatctggtctatgaaagacaacacttaacctatcccatctccagagaacatctcttca	2514
Db	2803	acaaatctggtctatgaaagacaacacttaacctatcccatctccagagaacatctcttca	2862
QY	2515	ttctcgtatgaaaaaccagagctctatgtgattcttgggtgtccacaactcagaactctcgaa	2574
Db	2863	ttctcgtatgaaaaaccagagctctatgtgattcttgggtgtccacaactcagaactctcgaa	2922
QY	2575	gagcgatcagacgccttaactgtgaagctttctagtctgtgtgtgtgtgtgtgtgtgtgtgt	2634
Db	2923	gagcgatcagacgccttaactgtgaagctttctagtctgtgtgtgtgtgtgtgtgtgtgtgt	2982
QY	2635	agagcagcttatgaaatatttcagcaactctgtctgtagtaaaaaacaatgtccattgaacaa	2694
Db	2983	agagcagcttatgaaatatttcagcaactctgtctgtagtaaaaaacaatgtccattgaacaa	3042
QY	2695	gaagctctccccaagaattcaagacacctagcaactagcgcaaaagcaatttaagccacc	2754
Db	3043	gaagctctccccaagaattcaagacacctagcaactagcgcaaaagcaatttaagccacc	3060
QY	2755	caccagctcttgaagcgcacatcaacgaggaataactcgtatactcttcaagtccagatcaag	2814
Db	3061	caccagctcttgaagcgcacatcaacgaggaataactcgtatactcttcaagtccagatcaag	3120
QY	2815	aggaatattgataatgatatccatactcagcttgtaaatgtgaagaaggaagattttgacatt	2874
Db	3121	aggaatattgataatgatatccatactcagcttgtaaatgtgaagaaggaagattttgacatt	3180
QY	2875	atgatactgagtgataaatccagaagcccccggaagcttccaaagaaacagaacactattta	2934
Db	3181	atgatactgagtgataaatccagaagcccccggaagcttccaaagaaacagaacactattta	3240
QY	2935	ttgtctcagatgtgagagagctctcgtgattatctggagatgtagatgtccccacatgttctaa	2994
Db	3241	ttgtctcagatgtgagagagctctcgtgattatctggagatgtagatgtccccacatgttctaa	3300
QY	2995	acagagctcagatgtgagagatgtccctccagcttcaagaagaatgttttccaagaatttactg	3054
Db	3301	acagagctcagatgtgagagatgtccctccagcttcaagaagaatgttttccaagaatttactg	3360

QY	3055	atggtcctcttaccacgaccttaccgctgaggaaactaaatgaacatttggacctcgg	3114
Db	3361	atggctcttaccacgaccttaccgctgaggaaactaaatgaacatttggacctcgg	3420
QY	3115	ggccatataaggcgaagttgaagataatctcatgttaacttcgaactgaagacct	3174
Db	3421	ggccatataaggcgaagttgaagataatctcatgttaacttcgaactgaagacct	3480
QY	3175	ctcgtccctatctctctatcttaccgttattcttataggaagaatcagaggcaagag	3234
Db	3481	ctcgtccctatctctctatcttaccgttattcttataggaagaatcagaggcaagag	3540
QY	3235	cagaacctagaaaaaactttgtcgaagctaaatgaaccaaaacttatttggaaatgac	3294
Db	3541	cagaacctagaaaaaactttgtcgaagctaaatgaaccaaaacttatttggaaatgac	3600
QY	3295	aacatcatatgaccccccacaaagaatgaatttaccgcgaagcctgggtatttctctg	3354
Db	3601	aacatcatatgaccccccacaaagaatgaatttaccgcgaagcctgggtatttctctg	3660
QY	3355	atgttgacctgaaaaaagatgtcacctcgaagcttgaattgagacctctcgtctgcaca	3414
Db	3661	atgttgacctgaaaaaagatgtcacctcgaagccttgaattgagacctctcgtctgcaca	3720
QY	3415	ctaacaacatgaacctgtctcatgtggagacaaatgacagtagaagaatttgcctgttt	3474
Db	3721	ctaacaacatgaacctgtctcatgtggagacaaatgacagtagaagaatttgcctgttt	3780
QY	3475	tcacacatcttttgatgaaacccaagaagctggtacctctcgtgaatatatggaagaacttga	3534
Db	3781	tcacacatcttttgatgaaacccaagaagctggtacctctcgtgaatatatggaagaacttga	3840
QY	3535	gggtccctctgaatatcatcagaatgagaatcccaactttaaagaataatctgccttcag	3594
Db	3841	gggtccctctgaatatcatcagaatgagaatcccaactttaaagaataatctgccttcag	3900
QY	3595	caatloaalvgctctacataatgagatacacctacgttgattgaatvgctcgaagtagcaaga	3654
Db	3901	caatloaalvgctctacataatgagatacacctacgttgattgaatvgctcgaagtagcaaga	3960
QY	3655	ttcgatgttatctgtctagaatgaggcgaatgaaaacttcacatctatctattcattcag	3714
Db	3961	ttcgatgttatctgtctagaatgaggcgaatgaaaacttcacatctatctattcattcag	4020
QY	3715	gacatgttctacgttaccgaaaaaagagagatataaatatgacgttacaatctctatc	3774
Db	4021	gacatgttctacgttaccgaaaaaagagagatataaatatgacgttacaatctctatc	4080
QY	3775	cagggtgttttttgagacagtggaaatgtttacacatccaaaacttggaaatttggcgggtgagat	3834
Db	4081	cagggtgttttttgagacagtggaaatgtttacacatccaaaacttggaaatttggcgggtgagat	4140
QY	3835	ggccttatctggcggagcatctacatgctcgggatatgacacacttttctgtgtgtacagcaata	3894
Db	4141	ggccttatctggcggagcatctacatgctcgggatatgacacacttttctgtgtgtacagcaata	4200
QY	3895	agtttcgaagacccccctgggaattggtcttcggaaacaatttgaagatttccagatttaacgatt	3954
Db	4201	agtttcgaagacccccctgggaattggtcttcggaaacaatttgaagatttccagatttaacgatt	4260
QY	3955	caggacaataatgacagttgggcccccaaaacttggccagacatcatattatccggatcaatca	4014
Db	4261	caggacaataatgacagttgggcccccaaaacttggccagacatcatattatccggatcaatca	4320
QY	4015	atgccttgaggacacaaaggagaccttttcttgatacaagttggaatctgtttggcaccaatga	4074
Db	4321	atgccttgaggacacaaaggagaccttttcttgatacaagttggaatctgtttggcaccaatga	4380
QY	4075	ttatttcacggatcaaaaaccccaagggtggcccgatcgaagaattctccagacctctacatctctc	4134
Db	4381	ttatttcacggatcaaaaaccccaagggtggcccgatcgaagaattctccagacctctacatctctc	4440
QY	4135	agtttatacatatgtatagctcttgatvgggaaagaagtgtgcagaacttatccgaaatttcca	4194



Db	4441	agttatcatcagatagctcttgatggaagatggcagactatccggaaattcca	4500
Qy	4195	ctggaacctaatatggtctctttggcaatgttgatctcatctggataaacaatatctt	4254
Db	4501	ctggaacctaatatggtctctttggcaatgttgatctcatctggataaacaatatctt	4560
Qy	4255	ttaacctccaatatatgtcgcgtatcacgcgtttgcaccccaacctatataagcttcgca	4314
Db	4561	ttaacctccaatatatgtcgcgtatcacgcgtttgcaccccaacctatataagcttcgca	4620
Qy	4315	gcactctcgcatgagatltgatggcctgtgattaaatagtctgcagatgcatltggaa	4374
Db	4621	gcactctcgcatgagatltgatggcctgtgattaaatagtctgcagatgcatltggaa	4680
Qy	4375	tggagagttaaagcaatatcagatgcaagatctgttcatcctactttaaccaatgtc	4434
Db	4681	tggagagttaaagcaatatcagatgcaagatctgttcatcctactttaaccaatgtc	4740
Qy	4435	ttgcacactgtctcctctcaaaaagctgacttcaacctccaaggaggatgaatgcctgca	4494
Db	4741	ttgcacactgtctcctctcaaaaagctgacttcaacctccaaggaggatgaatgcctgca	4800
Qy	4495	gacctcagtgtaataacccaagaagtgtgctgcaagtgtgacctccagaagaacatgaag	4554
Db	4801	gacctcagtgtaataacccaagaagtgtgctgcaagtgtgacctccagaagaacatgaag	4860
Qy	4555	tccacagagtaactatctcaaggagtaaatctcgtctaccagatgtatgttgaaggatc	4614
Db	4861	tccacagagtaactatctcaaggagtaaatctcgtctaccagatgtatgttgaaggatc	4920
Qy	4615	tccctactlccacagatccaagatgagccatagtggaactctcttttccaagaatggcaag	4674
Db	4921	tccctactlccacagatccaagatgagccatagtggaactctcttttccaagaatggcaag	4980
Qy	4675	ttaaggtttttcagggaataccaagactccttcaacctgtgtgtgaactctctgaccacc	4734
Db	4981	ttaaggtttttcagggaataccaagactccttcaacctgtgtgtgaactctctgaccacc	5040
Qy	4735	cgttactgactcgctacaccttcgaattcaccccccaagttgtgtcacacagatgcctga	4794
Db	5041	cgttactgactcgctacaccttcgaattcaccccccaagttgtgtcacacagatgcctga	5100
Qy	4795	ggatggaggtctctggcctcgagggcacaggacctctactga 4835	
Db	5101	ggatggaggtctctggcctcgagggcacaggacctctactga 5141	
RESULT 11			
AAN90654			
ID	AAN90654	standard; DNA; 4275 BP.	
AC	AAN90654;		
XX			
DT	26-JUN-1990	(first entry)	
XX			
DE	DNA encoding 740 Arg-1649 Glu human Factor VIII:C.		
XX			
KW	Human Factor VIII:C; Ad.RE.neo; 740 Arg-1649 Glu Factor VIII:C;		
KW	haemophilia A.		
XX			
OS	Homo sapiens.		
XX			
FN	Key	Location/Qualifiers	
FT	CDS	1..4275	
FT		/tag- a	
XX			
PN	EP306968-A.		
XX			
PD	15-MAR-1989.		
XX			
PF	09-SEP-1988;	88EP-0114769.	
XX			

PR 08-APR-1988; 88JP-0085454.  
XX  
PA (KAGA ) CHERO-SERO-THERAP.  
PA (TEIJ ) TEIJUN LTD.  
XX  
PI Sugiyama T, Masuda K, Tajima Y, Yonemura H;  
XX  
DR WPT; 1989-078467/11.  
DR P-PSDB; AAP91165.  
XX  
XX  
PT Prodn. of recombinant human Factor-VIII-C -  
PT using animal cells transformed with a vector contg. the gene for  
PT Factor VIII:C and a promoter  
XX  
XX  
PS Fig 1(1) - 1(13); ; 32pp; English.  
XX  
CC When translated, Arg-740 of the carboxyl terminus of the H chain is  
CC directly bonded by a peptide bond to Glu-1649 of the amino terminus of  
CC L chain. It is used to transform animal cells so that they produce  
CC human Factor VIII:C. A pref. expression vector is plasmid Ad.RE.neo.  
CC The expression vector has at least one promoter upstream of AAN90654.  
CC The transformants can constantly and continuously produce human Factor  
CC VIII:C in high yield on a commercial scale. The human Factor VIII:C so  
CC produced is considered to corresp. to the smallest species of active and  
CC intact Factor VIII:C molecules in the human blood plasma. It is useful  
CC for treating haemophilia A patients.  
XX  
XX Sequence 4275 BP; 1245 A; 941 C; 945 G; 1144 T; 0 other;

Query Match	52.5%	Score 4173	DB 10	Length 4275
Best Local Similarity	98.0%	Pred. No. 0		
Matches 4270	Conservative	0	Mismatches	5
			Indels	84
			Gaps	
QY 477	gcacacagaagaatactactcgtggtgcagatgagacgtctcatggtgactatatagtcaaaagtgtat	536		
DB 1	gcacacagaagaatactactcgtggtgcagatgagacgtctcatggtgactatatagtcaaaagtgtat	60		
QY 537	ctcgtgtagctgacctgtgtgagcgcaagattctcctctctagatgycgcaaaatctttccatc	596		
DB 61	ctcgtgtagctgacctgtgtgagcgcaagattctcctctctagatgycgcaaaatctttccatc	120		
QY 597	aacacctcagtcgtgttacaacaaaagcctcgttttagaatctcaagttcaacttttaac	656		
DB 121	aacacctcagtcgtgttacaacaaaagcctcgttttagaatctcaagttcaacttttaac	180		
QY 657	atcgcgttaagccaaagcgcaaccctgagatggctcgtcagaagctccatcacatccaggtcaggt	716		
DB 181	atcgcgttaagccaaagcgcaaccctgagatggctcgtcagaagctccatcacatccaggtcaggt	240		
QY 717	tatgatacaagtggtcattacttaacttaagaacatggtctcccatcctgtcagttccatgct	776		
DB 241	tatgatacaagtggtcattacttaacttaagaacatggtctcccatcctgtcagttccatgct	300		
QY 777	gtctgtgtatctacttaattggaagctctctgaggagcctggaatgatgatgaacagcagttca	836		
DB 301	gtctgtgtatctacttaattggaagctctctgaggagcctggaatgatgatgaacagcagttca	360		
QY 837	aggaggaagaagaatgatataaagtctctccctgtgtggaagccatataatgtctcggaggtc	896		
DB 361	aggaggaagaagaatgatataaagtctctccctgtgtggaagccatataatgtctcggaggtc	420		
QY 897	ctgaaagaagaatggtctccaaatggtgcctctgaaccacatggtgccttaactactatattctct	956		
DB 421	ctgaaagaagaatggtctccaaatggtgcctctgaaccacatggtgccttaactactatattctct	480		
QY 957	catgtgagacctgtaaagaacttgatctcaagccctcatatggagccctactagtatagtga	1016		
DB 481	catgtgagacctgtgtaaagaacttgatctcaagccctcatatggagccctactagtatagtga	540		
QY 1017	gaaggagatctggccaaggaagaaagacacagacatcttgacaanaattatatactactttgct	1076		
DB 541	gaaggagatctggccaaggaagaaagacacacacatcttgacaanaattatatactacttttgc	600		



QY 1077 gtatttgatgaagggaagaagtgtgcactcagaaaacaaagaactccttgatcagatagg 1136  
|||||  
Db 601 gtatttgatgaagggaagaaglttgcaccacagaaacaaagaactcccttgcatagcagatagg 660  
QY 1137 gatgctgcatctgtcggcccttggtccttaaaaatgcacacagtcnaatgtgtatgttaaacagg 1196  
|||||  
Db 661 gatgctgcatctgtcggcccttggtccttaaaaatgcacacagtcnaatgtgtatgttaaacagg 720  
QY 1197 tctctgcagagtcctgatttgatgtgcacacaggaatcagtcctatgtgcatagttgatggaaatg 1256  
|||||  
Db 721 tctctacacaggtctgcatgtgatatgcacacaggaatcagtcctatgtgcatagttgatggaaatg 780  
QY 1257 ggaacacactcctggaagtcgaatcattctctcgaaggtcacacattctctgttgaggaaac 1316  
|||||  
Db 781 ggcacacactcctggaagtcgaatcattctctcgaaggtcacacattctctgttgaggaaac 840  
QY 1317 catgcgcagagcgtctcttggaatatctgcacaaataacttctctactgtctcaaacactcttg 1376  
|||||  
Db 841 catgcgcagagcgtctcttggaatatctgcacaaataacttctctactgtctcaaacactcttg 900  
QY 1377 atggaaccttgacacagtttctactgttttgtcatatctctctccacacaaatgtagcatg 1436  
|||||  
Db 901 atggaaccttgacacagtttctactgttttgtcatatctctctccacacaaatgtagcatg 960  
QY 1437 gaagcttatgtccaagtagagcagctgtctccagaggaagccccaactagcaatgtaaaaaaat 1496  
|||||  
Db 961 gaagcttatgtccaagtagagcagctgtctccagaggaagccccaactagcaatgtaaaaaaat 1020  
QY 1497 gaaagaacggaagaactatgatatgcatcttactatctctgaatgtaatgtgtgtagagttt 1556  
|||||  
Db 1021 gaagaagcggagaagactatgatatgcatcttactatctctgaatgtaatgtgtgtagagttt 1080  
QY 1557 gatgtatgaacactcctcctctctatccaatctgcgtcagttgtccaaagaagactcctaana 1616  
|||||  
Db 1081 gatgtatgaacactcctcctctctatccaatctgcgtcagttgtccaaagaagactcctaana 1140  
QY 1617 actgtgggtacatctatctgtcgtcgtgaagaggaagactgtgagactgtgccccttgctctc 1676  
|||||  
Db 1141 actgtgggtacatctatctgtcgtcgtcgaagaggaagactgtgagactgtgccccttgctctc 1200  
QY 1677 gcccccgaatgacagaagtltataaagaatcaatatttgaaacaaatgagccctcagcggattggt 1736  
|||||  
Db 1201 gcccccgaatgacagaagtltataaagaatcaatatttgaaacaaatgagccctcagcggattggt 1260  
QY 1737 aggaagaatcaaaaaagtcctgattatgtgcataacacagatgtaaacctttaaagccgtgaa 1796  
|||||  
Db 1261 aggaagaatcaaaaaagtcctgattatgtgcataacacagatgtaaacctttaaagccgtgaa 1320  
QY 1797 gctattcaagcatgatacaggaatcttggagacttactttaaaggaagtgtggaagacaca 1856  
|||||  
Db 1321 gctattcaagcatgatacaggaatcttggagacttactttaaaggaagtgtggaagacaca 1380  
QY 1857 cgtgttgatatttaagaatcaacgaacgagaccataataacatctaccctcaaggaatc 1916  
|||||  
Db 1381 cgtgttgatatttaagaatcaacgaacgagaccataataacatctaccctcaaggaatc 1440  
QY 1917 actgattcgcgtcctctgtattcaaggaatataccaaaaaggtgttaaaacatttgaaggaat 1976  
|||||  
Db 1441 actgattcgcgtcctctgtattcaaggaatataccaaaaaggtgttaaaacatttgaaggaat 1500  
QY 1977 ttcccaattcgcgcagaggaataatltcaaatataatggaagcagtcgtgaagaatggg 2036  
|||||  
Db 1501 ttcccaattcgcgcagaggaataatltcaaatataatggaagcagtcgtgaagaatggg 1560  
QY 2037 ccaactaatcagatctccgtgtgcctgacccgctattactctagtttcgttaaatatggag 2096  
|||||  
Db 1561 ccaactaatcagatctccgtgtgcctgacccgctattactctagtttcgttaaatatggag 1620  
QY 2097 agagaatctagttcaggaactaatgtgcacctcctcctcaatctgctacaagaagaatcgttagat 2156  
|||||  
Db 1621 agagaatctagttcaggaactaatgtgcacctcctcctcaatctgctacaagaagaatcgttagat 1680

QY 2157 caaagaggaacacagataatgtcacaagaaggaatgtcatcctgtttctgtatttgat 2216  
|||||  
Db 1681 caaagaggaacacagataatgtcacaagaaggaatgtcatcctgtttctgtatttgat 1740  
QY 2217 gagaacccgaagctgtgtacctacagaagaataatacaagcttctctcccaatccagcttga 2276  
|||||  
Db 1741 gagaacccgaagctgtgtacctacagaagaataatacaagcttctctcccaatccagcttga 1800  
QY 2277 gtgcagcttgagatccagagatltccaagcctccaacatcatgtcacagcataatgtgctat 2336  
|||||  
Db 1801 gtgcagcttgagatccagagatltccaagcctccaacatcatgtcacagcataatgtgctat 1860  
QY 2337 gtcttgatagtttgcaagtttgcagttgtgtgtgcataaggggtggaatactgttacaattcta 2396  
|||||  
Db 1861 gtcttgatagtttgcaagtttgcagttgtgtgtgtgcataaggggtggaatactgttacaattcta 1920  
QY 2397 agcatgtgagcacagactgtgacttcccttctgtctctctctctgtgatataccttcaaacac 2456  
|||||  
Db 1921 agcatgtgagcacagactgtgacttcccttctgtctctctctctgtgatataccttcaaacac 1980  
QY 2457 aaaaatgtctatgtaagaacacactcaacctatcccatctcaaggaagaactgtcttcatg 2516  
|||||  
Db 1981 aaaaatgtctatgtaagaacacactcaacctatcccatctcaaggaagaactgtcttcatg 2040  
QY 2517 tcgattggaaaacccaaggtctatgtgattctgggtgtccacaactcagacttccggaacaga 2576  
|||||  
Db 2041 tcgattggaaaacccaaggtctatgtgattctgggtgtccacaactcagacttccggaacaga 2100  
QY 2577 ggcattgaacgccttacttgaaagtttcttagttgtgtacaaagaacactgtgtatattacagag 2636  
|||||  
Db 2101 ggcattgaacgccttacttgaaagtttcttagttgtgtgtacaaagaacactgtgtatattacagag 2160  
QY 2637 gacagttatgtaagatatlttcagcactactgtctgtgagttaaacaaatgcatgtgaaccaaaga 2696  
|||||  
Db 2161 gacagttatgtaagatatlttcagcactactgtctgtgagttaaacaaatgcatgtgaaccaaaga 2220  
QY 2697 agcttctcccaagaattcaagcacaccttagcactagcgaaagaacatattatgtgcaaccca 2756  
|||||  
Db 2221 ----- 2220  
QY 2757 ccaaglttgaaacgcacacaaaggaataaactcgttactactctcagtcagatcaagaag 2816  
|||||  
Db 2221 -----gaataaactcgttactactctcagtcagatcaagaag 2256  
QY 2817 gaaattgactatgtatgataacatatcagttgtgaatgaaagaagaagatttgaactattat 2876  
|||||  
Db 2257 gaaattgactatgtatgataacatatcagttgtgaatgaaagaagaagatttgaactattat 2316  
QY 2877 gatgagatgtaaaatcagagcccccagcgttccaagaagaacacagcactatttatt 2936  
|||||  
Db 2317 gatgagatgtaaaatcagagcccccagcgttccaagaagaacacagcactatttatt 2376  
QY 2937 gctgcagtggaagagctctgggattatgggattgagtagctccacacatgttcttaagaac 2996  
|||||  
Db 2377 gctgcagtggaagagctctgggattatgggattgagtagctccacacatgttcttaagaac 2436  
QY 2997 agggctcagagtggtcagtgctccctcaggttcaagaagaagtgtgttcccaaggaattactgat 3056  
|||||  
Db 2437 agggctcagagtggtcagtgctccctcaggttcaagaagaagtgtgttcccaaggaattactgat 2496  
QY 3057 ggtcccttactcagcccttatacgtgtggaagaataatgtaaacatttggagcctctgggg 3116  
|||||  
Db 2497 ggtcccttactcagcccttatacgtgtggaagaataatgtaaacatttggagcctctgggg 2556  
QY 3117 ccatataaagagcagaagttgtaagaataataatcatgtgtaacttccagaaatccagccct 3176  
|||||  
Db 2557 ccatataaagagcagaagttgtaagaataataatcatgtgtaacttccagaaatccagccct 2616  
QY 3177 cgtccctattctcctattcttaagccttattctctatgaggaagatcagaagcagaagga 3236  
|||||  
Db 2617 cgtccctattctcctattcttaagccttattctctatgaggaagatcagaagcagaagga 2676  
QY 3237 gaacctagaaaaaacttctgtcaagcctaagtgaacccaaacttacttcttggaaggtgcaa 3296



|||||  
Db 2677 gaacctgaaaaaactctgtcaagcctaaatgaacccaactactcttctggaatgcaaa 2736  
QY 3297 catcatatgagcaccacacaaagatgattctgactgcaaaagccttggtctattctctgat 3356  
Db 2737 catcatatgagcaccacacaaagatgattctgactgcaaaagccttggtctattctctgat 2796  
QY 3357 gttagcctggaanaaagatgtagactcaagcctgatttggaacccctctgtgtctgcacact 3416  
Db 2797 gttagcctggaanaaagatgtagactcaagcctgatttggaacccctctgtgtctgcacact 2856  
QY 3417 aacacactgaacccctgtctcattggtgagacaagtgcatacagaagaattgtctgttttc 3476  
Db 2857 aacacactgaacccctgtctcattggtgagacaagtgcatacagaagaattgtctgttttc 2916  
QY 3477 accactcttgtaggaaccaaagcctgtgactcacttgaaataatgaaagaacgcgcag 3536  
Db 2917 accactcttgtaggaaccaaagcctgtgactcacttgaaataatgaaagaacgcgcag 2976  
QY 3537 gctccctgcaatataccagaatggaagatcccaactttaagaagaattatcgtcttcatagca 3596  
Db 2977 gctccctgcaatataccagaatggaagatcccaactttaagaagaattatcgtcttcatagca 3036  
QY 3597 atcaaatggtctacataatgatacactactggtcttagtaatgtgctccagatcaaaagatt 3656  
Db 3037 atcaaatggtctacataatgatacactactggtcttagtaatgtgctccagatcaaaagatt 3096  
QY 3657 cgaatggtactcgtcagacatggtgagcaaatgaaacatccactctcattcacttgatgga 3716  
Db 3097 cgaatggtactcgtcagacatggtgagcaaatgaaacatccactctcattcacttgatgga 3156  
QY 3717 catgtgtcactgttacaagaaaaaagaagatataaattgycactgtacactctatcca 3776  
Db 3157 catgtgtcactgttacaagaaaaaagaagatataaattgycactgtacactctatcca 3216  
QY 3777 ggtgttttttgagacagtgaataatgttaccatccaagctggaattgtggtggaatgac 3836  
Db 3217 ggtgttttttgagacagtgaataatgttaccatccaagctggaattgtggtggaatgac 3276  
QY 3837 cttaatggtgagcactctcagctgtggaatgagacaactttctgtgtgacagcaataag 3896  
Db 3277 cttaatggtgagcactctcagctgtggaatgagacaactttctgtgtgacagcaataag 3336  
QY 3897 tgtcagaactccctctggaatggtctctggaacatlagagatttccagattacagcttca 3956  
Db 3337 tgtcagaactccctctggaatggtctctggaacatlagagatttccagattacagcttca 3396  
QY 3957 ggaacataatgagacagtgtggtcccaaaagctgtgcagactctatctccggatccaatcaat 4016  
Db 3397 ggaacataatgagacagtgtggtcccaaaagctgtgcagactctatctccggatccaatcaat 3456  
QY 4017 gctctgagacacaaagaccccttctctggaatcaaggtgactgtgtgacacaaatgatt 4076  
Db 3457 gctctgagacacaaagaccccttctctggaatcaaggtgactgtgtgacacaaatgatt 3516  
QY 4077 attcaagcagcatcaagaccagaggtgtcccgctcagaagctctccagcctctacactccag 4136  
Db 3517 attcaagcagcatcaagaccagaggtgtcccgctcagaagctctccagcctctacactccag 3576  
QY 4137 tttaatcatagtaatgactgtgagtggaagaagtgtgcagactctacgaggaattccact 4196  
Db 3577 tttaatcatagtaatgactgtgagtggaagaagtgtgcagactctacgaggaattccact 3636  
QY 4197 ggaaccttaatggtctcttctggaatgtgatacactctggaataaacaacataatttt 4256  
Db 3637 ggaaccttaatggtctcttctggaatgtgatacactctggaataaacaacataatttt 3696  
QY 4257 aacctccaattatgtcgtgatacaccgttctgcacccaactcattatagcatlctgcagc 4316  
Db 3697 aacctccaattatgtcgtgatacaccgttctgcacccaactcattatagcatlctgcagc 3756  
QY 4317 acctctgcagatgagtgatgtggtgtgtgatttaaatgttgacagatgccattgtggaatg 4376  
|||||

Db 3757 acctctgcagatgagtgatgtggtgtgtgatttaaatgtgtcagcatgtccattgtggaatg 3816  
QY 4377 gagagttaaagcacaatactagatgcacagatctactgtctatccactctacttaccataatgttt 4436  
Db 3817 gagagttaaagcacaatactagatgcacagatctactgtctatccactctacttaccataatgttt 3876  
QY 4437 gccacctgtctctcttcaaaaagctgcacttcaacctccaagggaggaatgtaatgtcctgaga 4496  
Db 3877 gccacctgtctctcttcaaaaagctgcacttcaacctccaagggaggaatgtaatgtcctgaga 3936  
QY 4497 ccttcagggaataatccaaaagatggtgtgcagaagtgtgcagcttccagaagacataatgaatg 4556  
Db 3937 ccttcagggaataatccaaaagatggtgtgcagaagtgtgcagcttccagaagacataatgaatg 3996  
QY 4557 acaggaatgaactactcagaaggaataatctcgtctaccagaatgataatgaagatctc 4616  
Db 3997 acaggaatgaactactcagaagggagtaaatctcgtctaccagaatgataatgaagatctc 4056  
QY 4617 ctcatctccagcagtcacaagatggtccatcagtgtgactctcttttccaagaatgcaagatga 4676  
Db 4057 ctcatctccagcagtcacaagatggtccatcagtgtgactctcttttccaagaatgcaagatga 4116  
QY 4677 aaggttttccaagggaatacgaactccttcaacactgtgtggaactctctagaccacag 4736  
Db 4117 aaggttttccaagggaatacgaactccttcaacactgtgtggaactctctagaccacag 4176  
QY 4737 ttactgactcgtctactcttcgaattcaccocccagatgtgtgtgacacagatgtcccttagg 4796  
Db 4177 ttactgactcgtctactcttcgaattcaccocccagatgtgtgtgacacagatgtcccttagg 4236  
QY 4797 atggaagttctgtggtctgcgaggtcacaggaactctactga 4835  
Db 4237 atggaagttctgtggtctgcgaggtcacaggaactctactga 4275  
  
RESULT 12  
AAN80446  
ID AAN80446 standard; DNA; 4275 BP.  
XX  
AC AAN80446;  
XX  
XX 10-OCT-1990 (first entry)  
DT  
XX Modified factor VIII:C sequence with the R740-El649 deletion.  
DE  
XX  
XX Modified factor VIII:C; haemophilia; procoagulant;  
KW blood coagulation; RE deletion; ss.  
XX  
XX Homo sapiens.  
OS  
XX  
XX W08800831-A.  
PN  
XX  
XX 11-FEB-1988.  
PD  
XX  
XX 31-JUL-1987; 87WO-US01814.  
PF  
XX  
XX 01-AUG-1986; 86US-0893375.  
PR  
XX  
XX (BIOJ ) BIOGEN NV.  
PA  
XX  
XX Pasek MP;  
PI  
XX  
XX WPI: 1988-049866/07.  
DR  
XX P-PSDB: AAP80267.  
XX  
PT New DNA sequences encoding modified factor VIII:C - with deletion of DNA  
PT encoding maturation polypeptide, useful for high yield transformation.  
PS  
XX  
XX Claim 3; Page 44-45-46-47; 97pp; English.  
CC The entire sequence encoding the maturation polypeptide of  
CC factor VIII:C is deleted, i.e. Arg 740-Glu 1649.  
CC The full length Factor VIII:C cDNA has two changes with respect to the



CC published sequence (EPO application 160457):  
CC CTC to CTA at Leu 242 and TTC to CTC change at amino acid residue 1880  
CC (Phe to Leu). The product is produced in approx. 20 times higher  
CC yields than previous recombinant produced factor VIII:C and are more  
CC easily purified. The peptide is used for treating haemophilia A, both  
CC acute and prolonged bleeding.  
CC See also AAN8044 and AAN80447.  
CX

SQ Sequence 4275 BP; 1245 A; 940 C; 946 G; 1144 T; 0 other;

Query Match 52.5%; Score 4171.6; DB 9; Length 4275;  
Best Local Similarity 98.0%; Pred. No. 0;  
Matches 4268; Conservative 0; Mismatches 4; Indels 84; Gaps 1;

QY 477 gccacagaataactactcctggtgcagtggaactgtcatgtcaggtatcgaagaatgat 536  
DB 4 gccacagaataactactcctggtgcagtggaactgtcatgtcaggtatcgaagaatgat 63  
QY 537 ctccgttagctcctgtgtgagcaagatllccctctagagtgccaaaatcttccattc 596  
DB 64 ctccgttagctcctgtgtgagcaagatllccctctagagtgccaaaatcttccattc 123  
QY 597 aacacctcagtcgtgtacaaaagaactcgttvtgagaatcacggttcacctttcaac 656  
DB 124 aacacctcagtcgtgtacaaaagaactcgttvtgagaatcacggttcacctttcaac 183  
QY 657 atgcctagcccaaggccaccttgatggtctgttagtgcctaccatccaggcgaagtt 716  
DB 184 atgcctagcccaaggccaccttgatggtctgttagtgcctaccatccaggcgaagtt 243  
QY 717 tatgatacagtggtcatctacacttaagaacatgtgcttccatccctgtcagctcatagt 776  
DB 244 tatgatacagtggtcatctacacttaagaacatgtgcttccatccctgtcagctcatagt 303  
QY 777 gtgtgttatctcctcctgtgaaaagcttctcgaggagctgaatatgtatgtacagaacagttcaa 836  
DB 304 gtgtgttatctcctcctgtgaaaagcttctcgaggagctgaatatgtatgtacagaacagttcaa 363  
QY 837 agggagaagaagaatgtataaagcttccctgtgtggaagccatacatatgtctggcaggtc 896  
DB 364 agggagaagaagaatgtataaagcttccctgtgtggaagccatacatatgtctggcaggtc 423  
QY 897 ctgaagaagaatgttccaatgtgcctgtgacccactgtgccttactactacatcttctc 956  
DB 424 ctgaagaagaatgttccaatgtgcctgtgacccactgtgccttactactacatcttctc 483  
QY 957 catgtgacacctgtgaaaagacttgaaattcagggcctcatcttgagccctactagatgtaga 1016  
DB 484 catgtgacacctgtgaaaagacttgaaattcagggcctcatcttgagccctactagatgtaga 543  
QY 1017 gaaaggagctctgagcaaggaaagacacagacctgtgacaatltatactacttttgcct 1076  
DB 544 gaaaggagctctgagcaaggaaagacacacacctgtgacaatltatactacttttgcct 603  
QY 1077 gtatttgaagaaggaaaagtltggacctcagaaaacaaagaactcctgtatgcagatag 1136  
DB 604 gtatttgaagaaggaaaagtltggacctcagaaaacaaagaactcctgtatgcagatag 663  
QY 1137 gatgtctcatctgtctgggacctgtgctcaaaatgtcacacagtcnaatgttatgttaacagg 1196  
DB 664 gatgtctcatctgtctgggacctgtgctcaaaatgtcacacagtcnaatgttatgttaacagg 723  
QY 1197 tctctgcaagtgcttgatgtgatgccaagaaatcaatctcatctgtgcatgttgatagatg 1256  
DB 724 tctctcaacaagtgcttgatgtgatgccaagaaatcaatctcatctgtgcatgttgatagatg 783  
QY 1257 ggcacacacctcctgaagtgcactcaatatctcctcgaaggtcacacattctctgttgaagaac 1316  
DB 784 ggcacacacctcctgaagtgcactcaatatctcctcgaaggtcacacattctctgttgaagaac 843  
QY 1317 catgccaagcgctccttggaatatctgcacaataacttctccttactgtccaacactcttg 1376

DB 844 catgccaagcgctccttggaatatctgcacaataacttctccttactgtccaacactcttg 903  
QY 1377 atggaaccttggaacggttctctctgtttgttatatcttccaccacaagatgagctg 1436  
DB 904 atggaaccttggaacggttctctctgtttgttatatcttccaccacaagatgagctg 963  
QY 1437 gaaacttatgtcaaatgatacagcgctgcacagaggaaccccaactcagatgaaaatat 1496  
DB 964 gaaacttatgtcaaatgatacagcgctgcacagaggaaccccaactcagatgaaaatat 1023  
QY 1497 gaagaagcggagaagactatgatgatctctactgtatcttgaaatgtgatvtgtcaggttt 1556  
DB 1024 gaagaagcggagaagactatgatgatctctactgtatcttgaaatgtgatvtgtcaggttt 1083  
QY 1557 gatgatgcaactctcctctcttataccaagtctgactgaattggccaagaacatccctaa 1616  
DB 1084 gatgatgcaactctcctctcttataccaagtctgactgaattggccaagaacatccctaa 1143  
QY 1617 acttggttacattacatgtcgtgaaagaggagcttggaactatgtcccttaagctcctc 1676  
DB 1144 acttggttacattacatgtcgtgaaagaggagcttggaactatgtcccttaagctcctc 1203  
QY 1677 gcccccgatgacagaagttataaagaatcaatatltgaaacatgtgcctcagcggaattgt 1736  
DB 1204 gcccccgatgacagaagttataaagaatcaatatltgaaacatgtgcctcagcggaattgt 1263  
QY 1737 aggaagtacaaaagaagtcggattatgtgcacacagatgaaacctttaaagctgtgaa 1796  
DB 1264 aggaagtacaaaagaagtcggattatgtgcacacagatgaaacctttaaagctgtgaa 1323  
QY 1797 gctatccagcatgaaacaggaatcttggagaccttactcttctgtgggaagtgtgaaaccca 1856  
DB 1324 gctatccagcatgaaacaggaatcttggagaccttactcttctgtgggaagtgtgaaaccca 1383  
QY 1857 ctgtgttatataattaaagatccaagcaagacacataatacattaccctcacaaggaatc 1916  
DB 1384 ctgtgttatataattaaagatccaagcaagacacataatacattaccctcacaaggaatc 1443  
QY 1917 actgatgtccgtctctgttatcttcaagggagttaccaaaaggtgtaaaacatttgaaggt 1976  
DB 1444 actgatgtccgtctctgttatcttcaagggagttaccaaaaggtgtaaaacatttgaaggt 1503  
QY 1977 ttccaactctgcagaggaataattcaaalataatgtgaagatgtcgtagaagtgtg 2036  
DB 1504 ttccaactctgcagaggaataattcaaalataatgtgaagatgtcgtagaagtgtg 1563  
QY 2037 ccaactaaatcagatccctcggtgcctgtgacccgctattactcttagtttgcgttaatatgag 2096  
DB 1564 ccaactaaatcagatccctcggtgcctgtgacccgctattactcttagtttgcgttaatatgag 1623  
QY 2097 agaatctagcttcaagagctaatgtgcccctctccatcgtctcaagaagatctgtagat 2156  
DB 1624 agaatctagcttcaagagctaatgtgcccctctccatcgtctcaagaagatctgtagat 1683  
QY 2157 caaagaggaacacagataatgtcagacaagagaatgtcaccctgttttcgtaattgtat 2216  
DB 1684 caaagaggaacacagataatgtcagacaagagaatgtcaccctgttttcgtaattgtat 1743  
QY 2217 gagaacccgaagctgtgtacctacagagaataatacaacgcttctccccaatccagctgga 2276  
DB 1744 gagaacccgaagctgtgtacctacagagaataatacaacgcttctccccaatccagctgga 1803  
QY 2277 gtgcagcttgaggaatccagaattccaaggccccaactcatgtgacacagatcaatagggcat 2336  
DB 1804 gtgcagcttgaggaatccagaattccaaggccccaactcatgtgacacagatcaatagggcat 1863  
QY 2337 gtttttgatagttgtcaggtgtcagttgttctgtcatgaggtgtgcatatctgatactcta 2396  
DB 1864 gtttttgatagttgtcaggtgtcagttgttctgtcatgaggtgtgcatatctgatactcta 1923  
QY 2397 agcatgtgagacagagactgaactccttctgtctcttctcctgtgatataccttaaacac 2456  
DB 1924 agcatgtgagacagagactgaactccttctgtctcttctcctgtgatataccttaaacac 1983



QY	2457	aaaaatggtctataatgaagacaacatccacccattcccatctctcagaagaaactggtctctatg	2516
Db	1984	aaaaatggtctataatgaagacaacatccacccattcccatctctcagaagaaactggtctctatg	2043
QY	2517	tcgaatggaanaaccagaatctcatatgattcttggtgtgcacaaactcgaacttcgcgaacga	2576
Db	2044	tcgaatggaanaaccagaatctcatatgattcttggtgtgcacaaactcgaacttcgcgaacga	2103
QY	2577	ggcaatgacccgctctacttgaaaggttctctagtgtgacaaagaaacactggtgtattatcacag	2636
Db	2104	ggcaatgacccgctctacttgaaaggttctctagtgtgacaaagaaacactggtgtattatcacag	2163
QY	2637	gacagttatgaaagatatattcagaactctctgtgtgataaanaacatgcatcttgaaacaga	2696
Db	2164	gacagttatgaaagatatattcagaactctctgtgtgataaanaacatgcatcttgaaacaga	2223
QY	2697	agctctcccaagatatcaagaacacccctagcactagcgaaagacaaatttaatgcccacca	2756
Db	2224	agctctcccaagatatcaagaacacccctagcactagcgaaagacaaatttaatgcccacca	2223
QY	2757	ccaagtctgaaacgcgcacatcaacgvggaaaaatactcgtactactcttcagtcagatcaagag	2816
Db	2224	ccaagtctgaaacgcgcacatcaacgvggaaaaatactcgtactactcttcagtcagatcaagag	2259
QY	2817	gaaatgactatgatatataacatatatcagtgtgaaatgaaagaaagaaattttgacatttat	2876
Db	2260	gaaatgactatgatatataacatatatcagtgtgaaatgaaagaaagaaattttgacatttat	2319
QY	2877	gatgagatgaaataatcagaaccccgccagcttccaaagaanaacagacatatatttt	2936
Db	2320	gatgagatgaaataatcagaaccccgccagcttccaaagaanaacagacatatatttt	2379
QY	2937	gctgcagtggaagagctcgvvggattatgvggattgagtagctccccaacatggtcttaagaaac	2996
Db	2380	gctgcagtggaagagctcgvvggattatgvggattgagtagctccccaacatggtcttaagaaac	2439
QY	2997	agggccacagatgagcagtggtccctccagtttcagaagaattgttttcagaagaatttaactat	3056
Db	2440	agggccacagatgagcagtggtccctccagtttcagaagaattgttttcagaagaatttaactat	2499
QY	3057	ggctcccttaactcagcccttatataccgtgcggaactaaatgaaactttggagctccctggg	3116
Db	2500	ggctcccttaactcagcccttatataccgtgcggaactaaatgaaactttggagctccctggg	2559
QY	3117	ccatatataagacagagaagtgtgaagataataatactgtgtaacttccagaatccagccctc	3176
Db	2560	ccatatataagacagagaagtgtgaagataataatactgtgtaacttccagaatccagccctc	2619
QY	3177	cgctcccatctcccttcatatctcagaacttatctctatagagaagatccaaagcaagagaca	3236
Db	2620	cgctcccatctcccttcatatctcagaacttatctctatagagaagatccaaagcaagagaca	2679
QY	3237	gaacctagaaaaaaacttgttcaagccctaataatgaaacaaacttaacttcttgaaagtgc	3296
Db	2680	gaacctagaaaaaaacttgttcaagccctaataatgaaacaaacttaacttcttgaaagtgc	2739
QY	3297	catcatatgagccaccctaagaatgagtttgactgcaagagctgggcttatttcttgat	3356
Db	2740	catcatatgagccaccctaagaatgagtttgactgcaagagctgggcttatttcttgat	2799
QY	3357	gttgcacctgaaaaaagatgtgcactccagcctgcatgttgaaaccccttctgcttgccacact	3416
Db	2800	gttgcacctgaaaaaagatgtgcactccagcctgcatgttgaaaccccttctgcttgccacact	2859
QY	3417	aacacacatgaaaccctgctcatctgggagacaagtgcacagtaacagaatctgctglttttc	3476
Db	2860	aacacacatgaaaccctgctcatctgggagacaagtgcacagtaacagaatctgctglttttc	2919
QY	3477	accacatttgatgagcccaaaaagctggttactctcaacttgaaaaataatgtaagaagaactgcag	3536
Db	2920	accacatttgatgagcccaaaaagctggttactctcaacttgaaaaataatgtaagaagaactgcag	2979

QY	3537	gcccctgcgaatataccagatvgaaagatccaccattttaaagaatatacgtctccatgca	3536
Dh	2980	gctccctgcgaatataccagatvgaaagatccaccattttaaagaatatacgtctccatgca	3039
QY	3597	atcaatgctcaataaayggatatacctcgtgttaagtatgctccagatataaagagt	3656
Dh	3040	atcaatgctcaataaayggatatacctcgtgttaagtatgctccagatataaagagt	3099
QY	3657	cgatgctatcgtctccagatcagvggcaggaatvgaaataccatcttatacttcagtgga	3716
Dh	3100	cgatgctatcgtctccagatcagvggcaggaatvgaaataccatcttatacttcagtgga	3159
QY	3717	catgctgctactgtagcaaaaaaagagagatataaaygcactgtacatctccatcca	3776
Dh	3160	catgctgctactgtagcaaaaaaagagagatataaaygcactgtacatctccatcca	3219
QY	3777	ggctgttttttgagaagtggaagtgtacccatccaaagcgggaattgggggttggaatgc	3836
Dh	3220	ggctgttttttgagaagtggaagtgtacccatccaaagcgggaattgggggttggaatgc	3279
QY	3837	cttatgcgcagatccatcacatgctcggagatagaacacttttctgtgtacgcacataag	3896
Dh	3280	cttatgcgcagatccatcacatgctcggagatagaacacttttctgtgtacgcacataag	3339
QY	3897	tgccaagactccctcgtggaatgctctctgcacacatagaagattccaagattcaacgtcca	3956
Dh	3340	tgccaagactccctcgtggaatgctctctgcacacatagaagattccaagattcaacgtcca	3399
QY	3957	ggacaatatatggacaagtgggcccacaagctcggcagaactcatattccggtatcaataat	4016
Dh	3400	ggacaatatatggacaagtgggcccacaagctcggcagaactcatattccggtatcaataat	3459
QY	4017	gcccggagacacaagggagcccttctcttgatataaagtgtacgtcttgccacaataagt	4076
Dh	3460	gcccggagacacaagggagcccttctcttgatataaagtgtacgtcttgccacaataagt	3519
QY	4077	atccaagcatccaagaccagaagtgcccgctcagaagttctccaagcctcacaactctccag	4136
Dh	3520	atccaagcatccaagaccagaagtgcccgctcagaagttctccaagcctcacaactctccag	3579
QY	4137	ttttcatcatgataatcttgatcttggaagaatggcagaactatccggagaattccact	4196
Dh	3580	ttttcatcatgataatcttgatcttggaagaatggcagaactatccggagaattccact	3639
QY	4197	ggaaccttaatgycctcctcttgccaatgttgatctacgcgggataaaacacatactttt	4256
Dh	3640	ggaaccttaatgycctcctcttgccaatgttgatctacgcgggataaaacacatactttt	3699
QY	4257	aaacctccaatlatctgcatacatatccglttgaccaccaactcatatagaacttcgcagc	4316
Dh	3700	aaacctccaatlatctgcatacatatccglttgaccaccaactcatatagaacttcgcagc	3759
QY	4317	actcttcgcatggagtttgatgggctgttgatttaaatagtgtgaagatgscattgggaatg	4376
Dh	3760	actcttcgcatggagtttgatgggctgttgatttaaatagtgtgaagatgscattgggaatg	3819
QY	4377	gagagtaaaagcaatatcagaatgcacagatatactgcttcaaccttataccaatactgtt	4436
Dh	3820	gagagtaaaagcaatatcagaatgcacagatatactgcttcaaccttataccaatactgtt	3879
QY	4437	ggcacctggtctccttccaaaaagctcgactcacctccaaggagagatgaatgctctgaga	4496
Dh	3880	ggcacctggtctccttccaaaaagctcgactcacctccaaggagagatgaatgctctgaga	3939
QY	4497	cctcagtgtaataatatacaaaaagctggctggcaagtggacttccagaagaacatgaaatc	4556
Dh	3940	cctcagtgtaataatatacaaaaagctggctggcaagtggacttccagaagaacatgaaatc	3999
QY	4557	acaagagtaataataatcagggagatgaataactcctcttacagagatlatatgtgaagaatc	4616
Dh	4000	acaagagtaataataatcagggagatgaataactcctcttacagagatlatatgtgaagaatc	4059
QY	4617	ctcatctccaagcagttcaagaatggccatcagcttggaactccttctttccaagaatgccaagt	4676



|||||  
Db 4060 ctcattccagcagtcgaatgagccatcagtgagctctcttttcgaatgagcaagta 4119  
QY 4677 aaggttttcagggaatacaagactccttcacactgtgtgtaactctagaccacgcg 4736  
|||||  
Db 4120 aaggttttcagggaatacaagactccttcacactgtgtgtaactcttagaccacgcg 4179  
QY 4737 ttactgactcgtcactccttcgattccaccccgaggttggtgtgacccagattgacctgagg 4796  
|||||  
Db 4180 ttactgactcgtcactccttcgattccaccccgaggttggtgtgacccagattgacctgagg 4239  
QY 4797 atggaggtctcggcgtcgacgagcagagacctctac 4832  
|||||  
Db 4240 atggaggtctcggcgtcgacgagcagacctctac 4275  
  
RESULT 13  
AAN80447  
ID AAN80447 standard; DNA: 4272 BP.  
XX  
AC AAN80447;  
XX  
DT 10-OCR-1990 (first entry)  
XX  
DE Modified factor VIII:C sequence with the R740-D1658 deletion.  
XX  
KM Modified factor VIII:C; haemophilia; procoagulant;  
KM blood coagulation; RD deletion; ss.  
XX  
OS Homo sapiens.  
XX  
PN MO8800831-A.  
XX  
PD 11-FEB-1988.  
XX  
PF 31-JUL-1987; 87MO-US01814.  
XX  
PR 01-AUG-1986; 86US-0893375.  
XX  
PA (BIOJ ) BIOGEN NV.  
XX  
PI Pasek MP;  
XX  
DR WPI: 1988-049866/07.  
XX  
PT P-PSDB; AAP80268.  
XX  
PS New DNA sequences encoding modified factor VIII:C - with deletion of DNA  
encoding maturation polypeptide, useful for high yield transformation.  
XX  
PS Claim 3; Page 47-48-49-50; 97pp; English.  
XX  
CC The RD deletion removes the DNA from Ser 741 to Ser 1657. A major part  
of the sequence encoding the maturation polypeptide of  
CC factor VIII:C is deleted, i.e. Gln 744 - Asp 1567.  
CC The full length Factor VIII:C cDNA has two changes with respect to the  
CC published sequence (EBO application 160457):  
CC CNG to CNA at Leu 242 and TTC to CTC change at amino acid residue 1880  
CC (Phe to Leu). The product is produced in approx. 20 times higher  
CC yields than previous recombinant produced factor VIII:C and are more  
CC easily purified. The peptide is used for treating haemophilia A, both  
CC acute and prolonged bleeding.  
CC See also AAN80444 and AAN80446.  
XX  
SQ Sequence 4272 BP; 1243 A; 941 C; 944 G; 1144 T; 0 other;

QY 537 ctggtgagctgctgtggagcagaatttctctctagatgagccaaatctttcattc 596  
Db 61 ctggtgagctgctgtggagcagaatttctctctagatgagccaaatctttcattc 120  
QY 597 aacacccagtcgtgtacaaagaacctgtgtttaaatccaggttcaccttccaac 656  
|||||  
Db 121 aacacccagtcgtgtacaaagaacctgtgtttaaatccaggttcaccttccaac 180  
QY 657 atcgctaagccaagccaacctgtgattggtctgtctaggtctccatccatccagctgaggt 716  
Db 181 atcgctaagccaagccaacctgtgattggtctgtctaggtctccatccatccagctgaggt 240  
QY 717 tatgatcacgtgtgtcattacaccttaagaacctgtgtctccatccatccagctgaggt 776  
Db 241 tatgatcacgtgtgtcattacaccttaagaacctgtgtctccatccatccagctgaggt 300  
QY 777 gtgtgtatccctactgaaagctcttgagggagctgatatgtatgtacagaccgtcaa 836  
Db 301 gtgtgtatccctactgaaagctcttgagggagctgatatgtatgtacagaccgtcaa 360  
QY 837 agggagaagaagatgataaagctctccctgtgtgagccaatacatatgtctgcaggtc 896  
Db 361 agggagaagaagatgataaagctctccctgtgtgagccaatacatatgtctgcaggtc 420  
QY 897 ctgaaagagaatgtgtccaaatgtgacctgtgacctgtgtcttactactatcttct 956  
Db 421 ctgaaagagaatgtgtccaaatgtgacctgtgacctgtgtcttactactatcttct 480  
QY 957 catgtgacctgtgtgaaagacttgaattcaagctctcaattgagccctactgattgata 1016  
Db 481 catgtgacctgtgtgaaagacttgaattcaagctctcaattgagccctactgattgata 540  
QY 1017 gaaggagctctgccaaggaagaacacagacctgtgacaatttactacttctgtct 1076  
Db 541 gaaggagctctgccaaggaagaacacacacctgtgacaatttactacttctgtct 600  
QY 1077 gtattgtatgaaggaaaagtgtgacctctcagaacaaagaacctctgtatgagagatag 1136  
Db 601 gtattgtatgaaggaaaagtgtgacctctcagaacaaagaacctctgtatgagagatag 660  
QY 1137 gatgtcatctgtctcgccgctgacctgaataatgtcacacagtcacatgtgtatgataag 1196  
Db 661 gatgtcatctgtctcgccgctgacctgaataatgtcacacagtcacatgtgtatgataag 720  
QY 1197 tctctcgaagctctgattgtgacccacaggaatcagtcataatgtgacatgtatgata 1256  
Db 721 tctctcgaagctctgattgtgacccacaggaatcagtcataatgtgacatgtatgata 780  
QY 1257 ggcacacactcctgaagtcagctcaatatctctctgaaagtcacacattctgtgagaac 1316  
Db 781 ggcacacactcctgaagtcagctcaatatctctctgaaagtcacacattctgtgagaac 840  
QY 1317 catcgcaagcgtctcttggaatctgcacataactctcttactgtctcaaacctctgt 1376  
Db 841 catcgcaagcgtctcttggaatctgcacataactctcttactgtctcaaacctctgt 900  
QY 1377 atggaacttggaagttcttactgtttgtcattatctcttccacacacatgataatgacat 1436  
Db 901 atggaacttggaagttcttactgtttgtcattatctcttccacacacatgataatgacat 960  
QY 1437 gaagcttatgtcaaaagtagcagctgtccagaggaaccccaactacagaaatgaanaaat 1496  
Db 961 gaagcttatgtcaaaagtagcagctgtccagaggaaccccaactacagaaatgaanaaat 1020  
QY 1497 gaagaagcgaagactatgataatgtacatcagatctgaatgtgagtgagtgatcaggttc 1556  
Db 1021 gaagaagcgaagactatgataatgtacatcagatctgaatgtgagtgagtgatcaggttc 1080  
QY 1557 gatgatgacaactctctctctatccaaatctgactcaggttgccaagaagcactctaa 1616  
Db 1081 gatgatgacaactctctctctatccaaatctgactcaggttgccaagaagcactctaa 1140



QY	1617	acttgggttaactatcattcgttcgtcgtgaagaaggagacttgggactatgcttcccttaagtcctc	1676
Db	1141	acttgggttaactatcattcgttcgtcgtgaagaaggagacttgggactatgcttcccttaagtcctc	1200
QY	1677	gcccccgatgcacaagtataaagtcaatatatttgacaacatggtcccttaagcggatttgg	1736
Db	1201	gcccccgatgcacaagtataaagtcaatatatttgacaacatggtcccttaagcggatttgg	1260
QY	1737	aggaagtacaacaaaagttccgaatttatgtgcatacacacagatgaaaccttaagaactcgttgaa	1796
Db	1261	aggaagtacaacaaaagttccgaatttatgtgcatacacacagatgaaaccttaagaactcgttgaa	1320
QY	1797	gcatatcagcatatgataccagaactcttggagaccttaacttattggtgggaagtctggaacaca	1856
Db	1321	gcatatcagcatatgataccagaactcttggagaccttaacttattggtgggaagtctggaacaca	1380
QY	1857	ctgttgattatatttaagaataccaagcaagcagacacataaactatcacccttaagcgaatc	1916
Db	1381	ctgttgattatatttaagaataccaagcaagcagacacataaactatcacccttaagcgaatc	1440
QY	1917	actgatgtccgtcctctgttatctcaagagagatataccaagaagtgtlaaacaatttgagaat	1976
Db	1441	actgatgtccgtcctctgttatctcaagagagatataccaagaagtgtlaaacaatttgagaat	1500
QY	1977	tttcccaattctgcacagagaaatatctcaatatataatgacagatgactgtgaagaatctgg	2036
Db	1501	tttcccaattctgcacagagaaatatctcaatatataatgacagatgactgtgaagaatctgg	1560
QY	2037	ccaactaatatagaatccctcgtgtgcctgcacccgctattactgagtcttcgttaattgag	2096
Db	1561	ccaactaatatagaatccctcgtgtgcctgcacccgctattactgagtcttcgttaattgag	1620
QY	2097	agaagatcagaactcaaggagactatgtgcctctccctcatctgcttaacaaagaactctgaa	2156
Db	1621	agaagatcagaactcaaggagactatgtgcctctccctcatctgcttaacaaagaactctgaa	1680
QY	2157	caaaagagaaaccaagataatgttcagacaagaaggaaatgtaaccgttttctgttatattg	2216
Db	1681	caaaagagaaaccaagataatgttcagacaagaaggaaatgtaaccgttttctgttatattg	1740
QY	2217	gagaacccgaagctgtgtccctcacagagataatacaagcgtttctcccaatcccaactgtga	2276
Db	1741	gagaacccgaagctgtgtccctcacagagataatacaagcgtttctcccaatcccaactgtga	1800
QY	2277	gttcagactctgagataaccagagttccaaagctcccaacatcatgacagacataatgtgctat	2336
Db	1801	gttcagactctgagataaccagagttccaaagctcccaacatcatgacagacataatgtgctat	1860
QY	2337	gtttttgatatgttgcaggtctgacagtttgttctgtgcatagtgggtgcacactgtgataactcta	2396
Db	1861	gtttttgatatgttgcaggtctgacagtttgttctgtgcatagtgggtgcacactgtgataactcta	1920
QY	2397	agcaattgagacacagactgacttccctctgtctctctccctcgcgataatacctcaaac	2456
Db	1921	agcaattgagacacagactgacttccctctgtctctctccctcgcgataatacctcaaac	1980
QY	2457	aaaaatgttctatgataagacacatccacatcccatctcccatctcaggaagaaactgttctcat	2516
Db	1981	aaaaatgttctatgataagacacatccacatcccatctcccatctcaggaagaaactgttctcat	2040
QY	2517	tcgatatgaaaaaccaggtctatatgatcttcgggtgtgcacaactcagacttctcgaaacaga	2576
Db	2041	tcgatatgaaaaaccaggtctatatgatcttcgggtgtgcacaactcagacttctcgaaacaga	2100
QY	2577	ggcatgacccgctctactacgaaggttttctagtgtgtgacaaagaacacgtgtgatatatacag	2636
Db	2101	ggcatgacccgctctactacgaaggttttctagtgtgtgacaaagaacacgtgtgatatatacag	2160
QY	2637	gacagttatgaagaataattcagcactactgtctgagttaaaaaaatgycatctgaaccaaaga	2696
Db	2161	gacagttatgaagaataattcagcactactgtctgagttaaaaaaatgycatctgaaccaaaga	2220
QY	2697	agcttctcccaagatctcaagcacccctcgtgcactagtcacaaagaacttaattatgcccacca	2756

[illegible]



```
Db 3217 ggtgttttgagacagtgaatgttacatccaagctggaatttggcgggtgaaatgc 3276
Oy 3837 cttaattggcagacatcacatgctgggatgagccactttttcgggtctacagaataag 3896
Db 3277 cttaattggcagacatcacatgctgggatgagccactttttcgggtctacagaataag 3336
Oy 3897 tgtacagctcccccgggaatggtctctgacacatagagaatttcagatcacgcttca 3956
Db 3337 tgtacagctcccccgggaatggtctctgacacatagagaatttcagatcacgcttca 3396
Oy 3957 ggcacatatgacagtgggcccaaaagctggccagacttcaatcctccgataccaat 4016
Db 3397 ggcacatatgacagtgggcccaaaagctggccagacttcaatcctccgataccaat 3456
Oy 4017 ggcctggagcccaagagagccctttcttgatcaaggatgatactgttggccaatgatc 4076
Db 3457 ggcctggagcccaagagagccctttcttgatcaaggatgatactgttggccaatgatc 3516
Oy 4077 attcagcgcatacgaagccagaggtgcccgtcaagaagttctccagcctctacatctcag 4136
Db 3517 attcagcgcatacgaagccagaggtgcccgtcaagaagttctccagcctctacatctcag 3576
Oy 4137 ttatcatcatgatactgtcttgatgggaagaagtggcagacttaccgaagaaattccact 4196
Db 3577 ttatcatcatgatactgtcttgatgggaagaagtggcagacttaccgaagaaattccact 3636
Oy 4197 ggaacctaatagtctctctttgtgcaatgtagattcatctgggataaaacaacaattttc 4256
Db 3637 ggaacctaatagtctctctttgtgcaatgtagattcatctgggataaaacaacaattttc 3696
Oy 4257 aacctccaattatgtctgcatacatccgtttgcaccaactcaatlaagacatcgagc 4316
Db 3697 aacctccaattatgtctgcatacatccgtttgcaccaactcaatlaagacatcgagc 3756
Oy 4317 acctctgcagtgaatgtgtagtggtgctgtgatttaaatgttgcagcatgccaatggaaatg 4376
Db 3757 acctctgcagtgaatgtgtagtggtgctgtgatttaaatgttgcagcatgccaatggaaatg 3816
Oy 4377 gagaagtaagaacataatagatgacagattactgctcatcacttaccataatatttc 4436
Db 3817 gagaagtaagaacataatagatgacagattactgctcatcacttaccataatatttc 3876
Oy 4437 gccacgtgctcctctcaaaaagctcgacttcaacctccaagggagtaatgctctggaga 4496
Db 3877 gccacgtgctcctctcaaaaagctcgacttcaacctccaagggagtaatgctctggaga 3936
Oy 4497 cctcaggtgataatccaagaaagtggctgcaagtggacttccagaagacaatgaaatgc 4556
Db 3937 cctcaggtgataatccaagaaagtggctgcaagtggacttccagaagacaatgaaatgc 3996
Oy 4557 accaagaataactactcaaggagtaaaatctctgttccagcagatgtaatggaaggagttc 4616
Db 3997 accaagaataactactcaaggagtaaaatctctgttccagcagatgtaatggaaggagttc 4056
Oy 4617 ctaatctccagcagctcaagatgagccatcaagtgagctctcttttttcaagaatgcaagta 4676
Db 4057 ctaatctccagcagctcaagatgagccatcaagtgagctctcttttttcaagaatgcaagta 4116
Oy 4677 aagggttttcaaggaaatcaagactccttcaacactgtgtgtaactctcagaagccacgcg 4736
Db 4117 aagggttttcaaggaaatcaagactccttcaacactgtgtgtaactctcagaagccacgcg 4176
Oy 4737 ttactgactgctacacttgaattcaaccccgagatggtgtgacagagattgcccftagg 4796
Db 4177 ttactgactgctacacttgaattcaaccccgagatggtgtgacagagattgcccftagg 4236
Oy 4797 atggaagttctggcctgagagcagacagacctctac 4832
Db 4237 atggaagttctggcctgagagcagacagacctctac 4272
```

RESULT 14  
AAx82258

```
ID AAX82258 standard; cDNA; 4373 BP.
XX AC
XX AAX82258;
XX DT
XX 18-AUG-1999 (first entry)
XX DE
XX Beta-domain deleted Factor VIII protein encoding cDNA.
XX KW
XX Factor VIII protein; gene modification; gene therapy; clinical disorder;
XX KW
XX splicing pattern; RNA processing; gene regulation; beta-domain; human;
XX OS
XX Homo sapiens.
XX PN
XX M09929848-A1.
XX PD
XX 17-JUN-1999.
XX PF
XX 25-NOV-1998; 98MO-US25354.
XX PR
XX 16-JAN-1998; 98US-0071596.
XX PR
XX 05-DEC-1997; 97US-0067614.
XX PA
XX (IMMU-) IMMUNE RESPONSE CORP.
XX PI
XX BidlIngmaier S, Gonzales JEN, Ill CR, Yang CO;
XX WPI; 1999-385602/32.
XX DR
XX P-PsDB; AAY21675.
XX PT
XX Genes and vectors exhibiting increased expression and novel splicing
XX PT
XX patterns; useful for expression of, e.g. beta-domain deleted factor
XX VIII
XX
XX Claim 9; Page 72-78; 123pp; English.
XX PS
XX
XX CC
XX The invention describes novel genes and vectors exhibiting increased
XX CC
XX expression and novel splicing patterns. It provides a gene encoding a
XX CC
XX Factor VIII protein, that comprises one or more consensus or near
XX CC
XX consensus splice sites which have been corrected to increase expression.
XX CC
XX The method, DNA sequences and expression vectors can be used to increase
XX CC
XX the expression of a gene, especially a Factor VIII gene. Genes containing
XX CC
XX modified 5' and/or 3' untranslated regions have optimized expression
XX CC
XX levels and tissue-specific expression. The methods are used for
XX CC
XX identification and correction of consensus splice sites; addition of
XX CC
XX introns; optimization of 5' and 3' untranslated regions and increase in
XX CC
XX cytoplasmic RNA accumulation. Hence the DNAs are useful in gene therapy
XX CC
XX to treat a clinical disorder, to study RNA processing and/or gene
XX CC
XX regulation. The present sequence represents a cDNA encoding a beta-domain
XX CC
XX deleted factor VIII protein.
XX SQ
XX Sequence 4373 BP; 1278 A; 1011 C; 938 G; 1146 T; 0 other;

Query Match 52.4%; Score 4161; DB 20; Length 4373;
Best Local Similarity 96.8%; Pred. No. 0;
Matches 4273; Conservative 0; Mismatches 100; Indels 42; Gaps 1;

Oy 420 atgcaatagagctcaccctctctctctgtgctcttgatcgatctgttagtgc 479
Db 1 atgcaatagagctcaccctctctctctgtgctcttgatcgatctgttagtgc 60

Oy 480 accaagaataactaactgggtgacagtggaactgtcaaggacatatagcaatgagatc 539
Db 61 accaagaataactaactgggtgacagtggaactgtcaaggacatatagcaatgagatc 120

Oy 540 ggtgagctgctgtgacgcaagaattccctcctagaagtgccaaatctttcatcacaac 599
Db 121 ggtgagctgctgtgtgacgcaagaattccctcctcgcgtgccaatctttcatcacaac 180

Oy 600 acctcagctgtgtacaaaagactctgttgtgaattcaaggttcaacctttcaacatc 659
Db 181 acctcagctgtgtacaaaagactctgttgtgaattcaaggttcaacctttcaacatc 240
```



[illegible]

QY	1740	aagracacaaaagccgattatctgacataacagatgaaacctttaaagctctgtaaac	1799
Db	1321	aagcacacaaaagccgattatctgacataacagatgaaacctttaaagctctgtaaac	1380
QY	1800	attcaagatgtaataagaaatctctggagcccttactttaaagggaagtctgagacacatg	1859
Db	1381	attcagcatgtaataagaaatctctggagcccttactttaaagggaagtctgagacacatg	1440
QY	1860	ttgattatatttaagaatccaagcaagcgcacataaacatctacccctaagaaatcact	1919
Db	1441	ctctatattatttaagaatccaagcaagcgcacataaacatctacccctaagaaatcacc	1500
QY	1920	gatctccgctctcttgatctcaagagagatccacaaggctgataaacatttgaagatttc	1979
Db	1501	gatctccgctctcttgatctcaagcagatccacaaggagtaaacatttgaagatttc	1560
QY	1980	ccaattctgcacagagaaataatccaataataatgacagctgctgtagaagaatgagcca	2039
Db	1561	ccaattctgcacagagaaataatccaataataatgacagctgctgtagaagaatgagcca	1620
QY	2040	actataacagatccctcggtgcgcgcgacccgctatactcagtttgcgttaatatggagga	2099
Db	1621	actataacagatccctcggtgcgcgcgacccgctatactcagtttgcgttaatatggagga	1680
QY	2100	gatactagcttcaagatcattctggccctctctcatctgctatacaagaagactgttagataca	2159
Db	1681	gatactagcttcaagatcattctggccctctctcatctgctatacaagaagactgttagataca	1740
QY	2160	agagagaacacagataatgctcagaacagaggaatgctacccgtttctctgtaatttgaatgag	2219
Db	1741	agagagaacacagataatgctcagaacagaggaatgctacccgtttctctgtaatttgaatgag	1800
QY	2220	aaccgaagctgtgataccctcacaaggaatatacaacgcttctccccaatccaaactgtgaagtg	2279
Db	1801	aaccgaagctgtgataccctcacaaggaatatacaacgcttctccccaatccaaactgtgaagtg	1860
QY	2280	caagcttgagagatcccaagagttccaagcctccaacatcaatgacacgaatcaatgctgatagtt	2339
Db	1861	caagcttgagagatcccaagagttccaagcctccaacatcaatgacacgaatcaatgctgatagtt	1920
QY	2340	tttgatagtttgcaggttgtaagttttcttgacagagtgtagcactgtagatctttaaagc	2399
Db	1921	ttgcagatgcttgcaggttgtaagttttcttgacagagtgtagcactgtagatcttctaagc	1980
QY	2400	attgagacacagaaatgtaactctctctctgctcttcttcctccggatataccttcaacacaaa	2459
Db	1981	attgagacacagaaatgtaactctctctctgctcttcttcctccggatataccttcaacacaaa	2040
QY	2460	atggtctatgtagagacaacatccaccattcccaattctcaggggaagctgtcttctatgctg	2519
Db	2041	atggtctatgtagagacaacatccaccattcccaattctcaggggaagctgtcttctatgctg	2100
QY	2520	atgtaaaaaacccaagttctatgtaattcttctgggtgtagcaaacatcagacttcttgagacagagagc	2579
Db	2101	atgtaaaaaacccaagactatgtaattcttctgggtgtagcaaacatcagacttcttgagacagagagc	2160
QY	2580	atgacgcgccttactgtaaggtttctctagttgtgacaagaacacactggtgatattatacgaagac	2639
Db	2161	atgacgcgccttactgtaaggtttctctcaagttgtgacaagaacacactggtgatattatacgaagac	2220
QY	2640	agttatgtaagataatttcagataactctgctgcgtagtaaaaaacaatgacatgtaaaccaagaagc	2699
Db	2221	agttatgtaagataatttcagataactctgctgcgtagtaaaaaacaatgacatgtaaaccaagaagc	2280
QY	2700	ttctccacagatccaagaacaccccttagcactagtcnaaagaacatttaatgcccaccaccaca	2759
Db	2281	ttctccacagatccaagaacaccccttagcactagtcnaaagaacatttaatgcccaccaccaca	2298
QY	2760	gtctctgtaaacgcgcatacaacggggaataaactcgttactactcttcagatcaagataaggaag	2819
Db	2399	gtctctgtaaacgcgcatacaacggggaataaactcgttactactcttcagatcaagataaggaag	2358
QY	2820	atttactatataataataacataataacttgaataaataaataaagaagattttgaacattatgat	2879



Db 2359 attgactatgatgtacatcatcatcaagtgtgaatgaagaagaattctgcacattatgat 2418  
QY 2880 gagatgaataatcagagcccccgcagccttcaaaaagaacaacgacattttatgtc 2939  
Db 2419 gaggtatgaataatcagagcccccgcagccttcaaaaagaacaacgacattttatgtc 2478  
QY 2940 gcaagtggagagcctctggatattatggatctggatctcccaacgattcttaagaacagg 2999  
Db 2479 gcaagtggagagcctctggatattatggatctggatctcccaacgattcttaagaacagg 2538  
QY 3000 gctcagagctggcagctgtccctcagctcagtaagaagtgtgttcccaagaaatttactatgtgc 3059  
Db 2539 gctcagagctggcagctgtccctcagctcagtaagaagtgtgttcccaagaaatttactatgtgc 2598  
QY 3060 tcccttactcagcccttatccctgtggagaactaaatgaacatttgggaactcctgggcca 3119  
Db 2599 tcccttactcagcccttatccctgtggagaactaaatgaacatttgggaactcctgggcca 2658  
QY 3120 tataaagaagcagaagttgaagataataatcatgtgaacttcagaatccagccctcgt 3179  
Db 2659 tataaagaagcagaagttgaagataataatcatgtgaacttcagaatccagccctcgt 2718  
QY 3180 cccatttcctctatcttctagccttatcttctatgaagaaga tcaagaggcaagagcagaa 3239  
Db 2719 cccatttcctctatcttccctcatcatatcatatgagaga tcaagaggcaagagcagaa 2778  
QY 3240 ccttagaaaaaacttctgtcaagccttaatgaacaacaaacttcttggaaagtgcacacat 3299  
Db 2779 ccttagaaaaaacttctgtcaagccttaatgaacaacaaacttcttggaaagtgcacacat 2838  
QY 3300 catactgacaccacttaagaatgaagtcttgacgtcaaacgctggccttatcttcgcagt 3359  
Db 2839 catactgacaccacttaagaatgaagtcttgacgtcaaacgctggccttatcttcgcagt 2898  
QY 3360 gaacctggaaaaaagatgtgctactcagcctgtatctggacccctctgtctgcacaactaac 3419  
Db 2899 gaacctggaaaaaagatgtgctactcagcctgtatctggacccctctgtctgcacaactaac 2958  
QY 3420 acactgaaacccctgtctcatctggagagacaagtgcacagtacacaggaatttgcctgttttcacc 3479  
Db 2959 acactgaaacccctgtctcatctggagagacaagtgcacagtacacaggaatttgcctgttttcacc 3018  
QY 3480 atcttctgtatgacccaagaagtctgtactctacgtcaaaatatacgtgaagaagaactgcaggcct 3539  
Db 3019 atcttctgtatgacccaagaagtctgtactctacgtcaaaatatacgtgaagaagaactgcaggcct 3078  
QY 3540 cccctgcaaatatccagatggaagatcccaactttaagaagaattatcgcttccatgcaatc 3599  
Db 3079 cccctgcaaatatccagatggaagatcccaactttaagaagaattatcgcttccatgcaatc 3138  
QY 3600 aatgtgtacaataatggaatacactacactcgtgcttagtaatggtcagatcacaagaattctga 3659  
Db 3139 aatgtgtacaataatggaatacactacactcgtgcttagtaatggtcagatcacaagaattctga 3198  
QY 3660 tggatctcgtcagcagatggtggagaatgaacaatccatctcatctatcttaagtgagcat 3719  
Db 3199 tggatctcgtcagcagatggtggagaatgaacaatccatctcatctatcttaagtgagcat 3258  
QY 3720 gtgttcaactgtacgaaaaaaagagagatataaaatggtcactgtacaatctcatccaggt 3779  
Db 3259 gtgttcaactgtacgaaaaaaagagagatataaaatggtcactgtacaatctcatccgga 3318  
QY 3780 gtttttgaagaagtggaatgtttacatcccaagctgaatcttggcgggtggaatgctcct 3839  
Db 3319 gtttttgaagaagtggaatgtttacatcccaagctgaatcttggcgggtggaatgctcct 3378  
QY 3840 attgcgagacatctacactgtcgtgagatgagcacactttctcgtgtgaacaagaatgtgt 3899  
Db 3379 attgcgagacatctacactgtcgtgagatgagcacactttctcgtgtgaacaagaatgtgt 3438  
QY 3900 cagaactccctgggaatggtctctgcacacattagagatttccagatccagcttcagga 3959  
|||||

Db 3439 cagactcccttgggaatggtctctgcacacatlagagatttccagattacagcttcagga 3498  
QY 3960 caatactgacagctgggcccccaagctggccagacttaattatccggatcaatcatgcc 4019  
Db 3499 caatactgacagctgggcccccaagctggccagacttaattatccggatcaatcatgcc 3558  
QY 4020 tggagaccagaagagcccttctctggatccaaggtggatctcgtttggaccacaatgatat 4079  
Db 3559 tggagaccagaagagcccttctctggatccaaggtggatctcgtttggaccacaatgatat 3618  
QY 4080 caggtcatcaagaagccagggtgcgccgtcagaagatctccagcctctacatctcactgt 4139  
Db 3619 caggtcatcaagaagccagggtgcgccgtcagaagatctccagcctctacatctcactgt 3678  
QY 4140 atcatcatgtatagtcttctgagtggaagaagtggcagacttatcagaagaattccactgga 4199  
Db 3679 atcatcatgtatagtcttctgagtggaagaagtggcagacttatcagaagaattccactgga 3738  
QY 4200 accttaatggtctctcttggcgaatgtgattcatcctggatataaaacacaatatcttttac 4259  
Db 3739 acctcctcagctctcttcttggcgaatgtgattcatcctggatataaaacacaatatcttttac 3798  
QY 4260 cctccaattatgtctcgatataatccgcttctgcaaccaactcatatagatctgcagcact 4319  
Db 3799 cctccaattatgtctcgatataatccgcttctgcaaccaactcatatagatctgcagcact 3858  
QY 4320 ctctgcagtgagtgatgtggtctgtgatttaaatagttgcagatgccatttgggaatggag 4379  
Db 3859 ctctgcagtgagtgatgtggtctgtgatttaaatagttgcagatgccatttgggaatggag 3918  
QY 4380 agtaagaacatatcagatgcagacagatctcgtctacatcccttacttacaatatgtttgcc 4439  
Db 3919 agtaagaacatatcagatgcagacagatctcgtctacatcccttacttacaatatgtttgcc 3978  
QY 4440 acctgtctccttccaaaagctcagactcactccaaaggagagatgaatgcttgcagacct 4499  
Db 3979 acctgtctccttccaaaagctcagactcactcacaagagagatgaatgcttgcagacct 4038  
QY 4500 caggtgaataatccaaaagatgtgctgtgaagtgtgacttccagaagacataatgaatgccaa 4559  
Db 4039 caggtgaataatccaaaagatgtgctgtgaagtgtgacttccagaagacataatgaatgccaa 4098  
QY 4560 ggaatgaactcagagggagtaaatctctgtcttaacagcatgtatgtgaagagtctctc 4619  
Db 4099 ggaatgaactcagagggagtaaatctctgtcttaacagcatgtatgtgaagagtctctc 4158  
QY 4620 atctccagcagtcagaatggtccatcagctgcctcttcttcagaatggtcacaagtgaag 4679  
Db 4159 atctccagcagtcagaatggtccatcagctgcctcttcttccaatggtcacaagtgaag 4218  
QY 4680 gtttttcaagggaataatcaagactccttcaacactgtgtggaactctctagaccacacgttta 4739  
Db 4219 gtttttcaagggaataatcaagactccttcaacactgtgtggaactctctagaccacacgttta 4278  
QY 4740 ctgactcgtcactccttgcgaatccaccccaagagtgtgggtgcacagatggcctctgagga 4799  
Db 4279 ctgactcgtcactccttgcgaatccaccccaagagtgtgggtgcacagatggcctctgagga 4338  
QY 4800 gaggttctggctgcgagagcacaggaactctactg 4834  
Db 4339 gaggttctggctgcgagagcacaggaactctactg 4373

RESULT 15  
AANB1544  
ID AANB1544 standard; DNA; 4830 BP.  
XX  
XX  
AANB1544;  
AC XX  
XX  
DT 04-DEC-1990 (first entry)  
XX  
XX Human Factor VIII-C analog having exon 14 deleted.  
XX



KM Human Factor VIII-C analog; exon deletion; coagulation disorders;  
KM haemophilia; ss.  
XX  
PN EP265778-A.  
XX  
PD 04-MAY-1988.  
XX  
PF 14-OCT-1987; 87EP-0115043.  
XX  
PR 15-OCT-1986; 86US-0919153.  
XX  
PA (RORE ) RORER INT OVERSEAS.  
XX  
PI Sarver N, Drohan W;  
XX  
DR WPI; 1988-120930/18.  
XX  
PT Human Factor VIII-C analogue free of other proteins - produced from  
XX recombinant DNA and used in the treatment of coagulation disorders.  
PS Claim 3; Page 13; 42pp; English.  
XX  
CC The genetically engineered analogue can provide a dependable and  
CC readily available therapeutic agent to be used in the treatment of  
CC haemophilia and coagulation disorders in humans or animals.  
CC See also AANB1543-45.  
XX  
XX  
SQ Sequence 4830 BP; 1392 A; 1091 C; 1058 G; 1289 T; 0 other;

Query Match 52.3%; Score 4152.2; DB 9; Length 4830;  
Best Local Similarity 94.8%; Pred. No. 0;  
Matches 4386; Conservative 0; Mismatches 43; Indels 198; Gaps 1;

QY 420 atgcaaatagagctctccacgtctcttctctgctcttgcagatctgcttgcgc 479  
DB 1 atgcaaatagagctctccacgtctcttctctgctcttgcagatctgcttgcgc 60  
QY 480 accgaagaatactacctggtgcaagtgcacgtcatggaactatgcaaatgctatc 539  
DB 61 accgaagaatactacctggtgcaagtgcacgtcatggaactatgcaaatgctatc 120  
QY 540 ggtgagctgctctgtagcgaagaattctccctagagtgccaaatctttccatcaac 599  
DB 121 ggtgagctgctctgtagcgaagaattctccctagagtgccaaatctttccatcaac 180  
QY 600 acccagtcggtgtaaaaaagactctgtttgtagaatccaggtccactttccaacac 659  
DB 181 acccagtcggtgtaaaaaagactctgtttgtagaatccaggtccactttccaacac 240  
QY 660 gctaagccaagccaacctgtagtggtctgtagtctcctaccatccaggtgaattat 719  
DB 241 gctaagccaagccaacctgtagtggtctgtagtctcctaccatccaggtgaattat 300  
QY 720 gatacagtggtcattacaacttaagaacatggtctccatcctgtcaagtcctatcgt 779  
DB 301 gatacagtggtcattacaacttaagaacatggtctccatcctgtcaagtcctatcgt 360  
QY 780 ggtgtatcctctctggaagctcttgtaggagctgtagatatagtatgacagaccgtaaaagg 839  
DB 361 ggtgtatcctctctggaagctcttgtaggagctgtagatatagtatgacagaccgtaaaagg 420  
QY 840 gagaagaagaatgataaagctctccctggtggaagccaataatgctcgaagtcctg 899  
DB 421 gagaagaagaatgataaagctctccctggtggaagccaataatgctcgaagtcctg 480  
QY 900 aagaagaatggtccaatggtcctctgacccactgtgcttaacctactatcttctcat 959  
DB 481 aagaagaatggtccaatggtcctctgacccactgtgcttaacctactatcttctcat 540  
QY 960 gtggaacgtgtaaaaagactgaattcaggtccatctgtagagccctactagttagagaa 1019  
DB 541 gtggaacgtgtaaaaagactgaattcaggtccatctgtagagccctactagttagagaa 600

QY 1020 gggagctgcccagaagaagacacagacctgtcacaaattatctacttctgtctga 1079  
DB 601 gggagctgcccagaagaagacacagacctgtcacaaattatctacttctgtctga 660  
QY 1080 ttgtatgaagggaagaaagtgtgacactcagaacaagaacctctgtatgcagataggat 1139  
DB 661 ttgtatgaagggaagaaagtgtgacactcagaacaagaacctctgtatgcagataggat 720  
QY 1140 gctgcatcgtcgtcggtcgtgcttaaaatgcacaagtcattgtttatgtaaaagttct 1199  
DB 721 gctgcatcgtcgtcggtcgtgcttaaaatgcacaagtcattgtttatgtaaaagttct 780  
QY 1200 ctgccaagctgattggaatgcccaggaatcgtctctgtgctgattggaatgggc 1259  
DB 781 ctgccaagctgattggaatgcccaggaatcgtctctgtgctgattggaatgggc 840  
QY 1260 accactcctgaagtgacactcaatattcctcgaaggtaacacattctgtggaagacct 1319  
DB 841 accactcctgaagtgacactcaatattcctcgaaggtaacacattctgtggaagacct 900  
QY 1320 cgcacggtcctcttggaatctcgcgaatacttcttactgtctcaaacctcttgatg 1379  
DB 901 cgcacggtcctcttggaatctcgcgaatacttcttactgtctcaaacctcttgatg 960  
QY 1380 gaccttgagaagttctctactgtttgtcatatctctccacacacatgatgatgaa 1439  
DB 961 gaccttgagaagttctctactgtttgtcatatctctccacacacatgatgatgaa 1020  
QY 1440 gcttatgtaaaagttagacagctgtctcagaagaaacccaactagaatgaaataatgaa 1499  
DB 1021 gcttatgtaaaagttagacagctgtctcagaagaaacccaactagaatgaaataatgaa 1080  
QY 1500 gaagcggaaagctatgtagatcttactatcttggaagtgagtggtgtaggtgat 1559  
DB 1081 gaagcggaaagctatgtagatcttactatcttggaagtgagtggtgtaggtgat 1140  
QY 1560 gatgacaactcctcctcttataccaatctgctgaagtgcgaagaagcatccaaact 1619  
DB 1141 gatgacaactcctcctcttataccaatctgctgaagtgcgaagaagcatccaaact 1200  
QY 1620 tgggtacattacattgctgcctcgaagaaggaactcgtgcctatctccttgcctcgc 1679  
DB 1201 tgggtacattacattgctgcctcgaagaaggaactcgtgcctatctccttgcctcgc 1260  
QY 1680 cccgattgacagaagttataaagctaatatttgaaacaatggccctcagcggattgtagg 1739  
DB 1261 cccgattgacagaagttataaagctaatatttgaaacaatggccctcagcggattgtagg 1320  
QY 1740 aagtaaaaaaagtcggaatttatggtacatacagaatgaacctttaagactgtgaagct 1799  
DB 1321 aagtaaaaaaagtcggaatttatggtacatacagaatgaacctttaagactgtgaagct 1380  
QY 1800 attcagcatgatacagaatctctggaaccttactttatggggaagtgtggaacaactg 1859  
DB 1381 attcagcatgatacagaatctctggaaccttactttatggggaagtgtggaacaactg 1440  
QY 1860 ttgattatatttaagaatccaagcagacacataatacatctaccctcagcgaataact 1919  
DB 1441 ttgattatatttaagaatccaagcagacacataatacatctaccctcagcgaataact 1500  
QY 1920 gatgtcgtctctctgtattcaaggaagattaccaaaagtgtgtaaacatttgaagatttt 1979  
DB 1501 gatgtcgtctctctgtattcaaggaagattaccaaaagtgtgtaaacatttgaagatttt 1560  
QY 1980 ccaattctgcaggaagaatatccaataataatgagcagtgactgttagaagaatgggcca 2039  
DB 1561 ccaattctgcaggaagaatatccaataataatgagcagtgactgttagaagaatgggcca 1620  
QY 2040 actaaatcagatcctcgtgtcctgtagccgctattactgtttcgtttaatatgagaga 2099  
DB 1621 actaaatcagatcctcgtgtcctgtagccgctattactgtttcgtttaatatgagaga 1680



QY	2100	gatacgtctcaggacattcgccctctccatctctgctacaaagaattcgtatgacaa	2159
Db	1681	gatacgtctcaggacattcgccctctccatctctgctacaaagaattcgtatgacaa	1740
QY	2160	agagaaacacagataatgctcagacaagaagaatgctcgtttcttgtaattgtagag	2219
Db	1741	agagaaacacagataatgctcagacaagaagaatgctcgtttcttgtaattgtagag	1800
QY	2220	aacggaagctgttacctacacagaaatataaagctttcccccataccgcgtgagtg	2279
Db	1801	aacggaagctgttacctacacagaaatataaagctttcccccataccgcgtgagtg	1860
QY	2280	cagcttggagatccagagttccaaagccctcacatcatgacagatcaagtgatgtt	2339
Db	1861	cagcttggagatccagagttccaaagccctcacatcatgacagatcaagtgatgtt	1920
QY	2340	tttgatagtttcgagttgctcagttgtttgcaatgagtgacatcgttatcttaagc	2399
Db	1921	tttgatagtttcgagttgctcagttgtttgcaatgagtgacatcgttatcttaagc	1980
QY	2400	attggagcacagactgacttctctgtcttctctctgataataccttcaaacacaa	2459
Db	1981	attggagcacagactgacttctctgtcttctctctgataataccttcaaacacaa	2040
QY	2460	atggtctatgaaacacacatcacccctattcccatctcaagagaagaatgcttcatg	2519
Db	2041	atggtctatgaaacacacatcacccctattcccatctcaagagaagaatgcttcatg	2100
QY	2520	atggaacaaacccaggtctctatggtctcgtggtgacacaaactcgtgaaacagagc	2579
Db	2101	atggaacaaacccaggtctctatggtctcgtggtgacacaaactcgtgaaacagagc	2160
QY	2580	atggccgcttactcgaaggtttctagtgtgtgacaagaacatggttattatgaagagc	2639
Db	2161	atggccgcttactcgaaggtttctagtgtgtgacaagaacatggttattatgaagagc	2220
QY	2640	agttatgaagatattcagcactactgctggtgaaacaaatgcatctgaaacaaagagc	2699
Db	2221	agttatgaagatattcagcactactgctggtgaaacaaatgcatctgaaacaaagagc	2280
QY	2700	ttctccacagaattca-----	2714
Db	2281	ttctccacagaattcaactattgatacctctgtcttgataatacacatagtactcagata	2340
QY	2715	-----	2714
Db	2341	ccaaagaagaagtgtgaaatcccaaagaagatccaccagaanaaacagcttttaagaaaaag	2400
QY	2715	-----	2714
Db	2401	gatacatttgcctcgtgaagcctgtgtgaaagcaatcatgcaatagacgaataatgag	2460
QY	2715	-----	2714
Db	2461	ggaacaaataagcccgaaatagaagtcacccgtgacagacaggttagagactgaaagcgtg	2520
QY	2742	tttaatgccccccacacagctcttgaaaagcacaatcaacgggaataaactgtactactct	2801
Db	2521	ttgcttcaaaacccacacagctcttgaaaagcacaatcaacgggaataaactgtactactct	2560
QY	2802	cagtcagatcaagaagaatgactatgataatgataccatlaagttgaaatgaaagaagaa	2861
Db	2581	cagtcagatcaagaagaatgactatgataatgataccatlaagttgaaatgaaagaagaa	2640
QY	2862	gattttgacattatgataatgataatgataatgataatgataatgataatgataatgata	2921
Db	2641	gattttgacattatgataatgataatgataatgataatgataatgataatgataatgata	2700
QY	2922	cgaacatttatttctgctgagtgagagagctgtgagattatgagatgagatgagatgagat	2981
Db	2701	cgaacatttatttctgctgagtgagagagctgtgagattatgagatgagatgagatgagat	2760
QY	2982	catgtcttaagaacaaaggtcagagtgagtgctccctcagttcaagaagaagtgttttc	3041

---

Db	2761	catgtcttaagaacaaaggtcagagtgagtgctccctcagttcaagaagaagtgttttc	2820
QY	3042	caggaatttactgatactgctccttactcgaagcccttataccgttgagaaactaaatgaaat	3101
Db	2821	caggaatttactgatactgctccttactcgaagcccttataccgttgagaaactaaatgaaat	2880
QY	3102	ttggagactccctgggacataataagaagtgagaatgtgaataatatacagtgtaacttc	3161
Db	2881	ttggagactccctgggacataataagaagtgagaatgtgaataatatacagtgtaacttc	2940
QY	3162	agaatcagagcctctcgtcccatcttctatctcgtacatcttcttatagaagaagat	3221
Db	2941	agaatcagagcctctcgtcccatcttctatctcgtacatcttcttatagaagaagat	3000
QY	3222	cagagcgaagagcagaacactagaaaaaactgttcaagcctaatagaaaccaaaacttac	3281
Db	3001	cagagcgaagagcagaacactagaaaaaactgttcaagcctaatagaaaccaaaacttac	3060
QY	3282	tttggaaagtgcacacatcatatgaccccaataagaatgagtttgatctgnaaagccctgg	3341
Db	3061	tttggaaagtgcacacatcatatgaccccaataagaatgagtttgatctgnaaagccctgg	3120
QY	3342	gcttattctctgatacttgacacttgaaaaagatgtgacatcagagcctgattgagacccct	3401
Db	3121	gcttattctctgatacttgacacttgaaaaagatgtgacatcagagcctgattgagacccct	3180
QY	3402	ctggtctgcacactaaacacatgaaacccctgtcactatgagagacaaagtgcacagaa	3461
Db	3181	ctggtctgcacactaaacacatgaaacccctgtcactatgagagacaaagtgcacagaa	3240
QY	3462	tttgctgtgttttcacacattttgtagagaccaaagcgtgtacttcactcgtgaaatatg	3521
Db	3241	tttgctgtgttttcacacattttgtagagaccaaagcgtgtacttcactcgtgaaatatg	3300
QY	3522	gaaagaacctgcagagcctccctgcaatatccagatggaagaatcccatctttaaagagaat	3581
Db	3301	gaaagaacctgcagagcctccctgcaatatccagatggaagaatcccatctttaaagagaat	3360
QY	3582	tatcgtcttcaatgcaatcaatgctacataataatgatacactactcgtgtatgagct	3641
Db	3361	tatcgtcttcaatgcaatcaatgctacataataatgatacactactcgtgtatgagct	3420
QY	3642	caggaatcaagagattcgatgatactgtcagacatggtgcagacatgaaataacatctct	3701
Db	3421	caggaatcaagagattcgatgatactgtcagacatggtgcagacatgaaataacatctct	3480
QY	3702	attcaattcagtgagacatgtgttctactgtacgaaaaaaagggagatlaaaatgagcagtg	3761
Db	3481	attcaattcagtgagacatgtgttctactgtacgaaaaaaagggagatlaaaatgagcagtg	3540
QY	3762	tacaattcttatccagatgtgttttggagaacagtgtgaaatgttaccatccaagcgtgaaatc	3821
Db	3541	tacaattcttatccagatgtgttttggagaacagtgtgaaatgttaccatccaagcgtgaaatc	3600
QY	3822	ttggcgggtggaaatgcttatttggcggagcactataatgctggaatggaacaaactttctg	3881
Db	3601	ttggcgggtggaaatgcttatttggcggagcactataatgctggaatggaacaaactttctg	3660
QY	3882	gtgtacagcaataaagtgctcagaccccttggaatgtgctctcgtgacacatataagatttc	3941
Db	3661	gtgtacagcaataaagtgctcagaccccttggaatgtgctctcgtgacacatataagatttc	3720
QY	3942	cagattacagcttcaaggaacatatgagcagtggtgcccacaaagctgtgcagactctat	4001
Db	3721	cagattacagcttcaaggaacatatgagcagtggtgcccacaaagctgtgcagactctat	3780
QY	4002	ttcggatcaatcaatgctcgtgagacacaaagagcccttctcgtgacaaagtgagatcgt	4061
Db	3781	ttcggatcaatcaatgctcgtgagacacaaagagcccttctcgtgacaaagtgagatcgt	3840
QY	4062	ttggaccaaattgattatccagcagatcaaaacccagagtgcccgctcagaagaagtctccagc	4121



Db 3841 ttggaccaatgatattatcacccgcatcaagaccagggtgccgtcagaagttctccagc 3900  
QY 4122 ctccacatctctcagtttatcatcatgtaagttctgtagtggaagaagtgagacttat 4181  
|||||  
Db 3901 ctccacatctctcagtttatcatcatgtaagttctgtagtggaagaagtgagacttat 3960  
QY 4182 cgaaggaattccactggaaccttaatgtctctcttggaatgtgatatcatctggata 4241  
|||||  
Db 3961 cgaaggaattccactggaaccttaatgtctctcttggaatgtgatatcatctggata 4020  
QY 4242 aaacacaatatcttaaccctccaattatgtctgatatcatccgtttgacccaactcat 4301  
|||||  
Db 4021 aaacacaatatcttaaccctccaattatgtctgatatcatccgtttgacccaactcat 4080  
QY 4302 tatagcatctcgacgacctcttcgcatgtagtvtgagtggtgtgatttaatatgttcagc 4361  
|||||  
Db 4081 tatagcatctcgacgacctcttcgcatgtagtvtgagtggtgtgatttaatatgttcagc 4140  
QY 4362 atgcaattgggaatggaagagtaaaagcaatatacagaatgcacagattactgtcatctac 4421  
|||||  
Db 4141 atgcaattgggaatggaagagtaaaagcaatatacagaatgcacagattactgtcatctac 4200  
QY 4422 tttaaccaataatgtttgcacctgtgtctctccaagaagctgcacttcaactccaaggagg 4481  
|||||  
Db 4201 tttaaccaataatgtttgcacctgtgtctctccaagaagctgcacttcaactccaaggagg 4260  
QY 4482 agtaatgccttggaagacctcagagtgaataatccaagaagtgctgcaagtgcacttccag 4541  
|||||  
Db 4261 agtaatgccttggaagacctcagagtgaataatccaagaagtgctgcaagtgcacttccag 4320  
QY 4542 aagacaatgaaagtcacagaggttaactactcaaggagtaaaatctctgttaccagcatg 4601  
|||||  
Db 4321 aagacaatgaaagtcacagaggttaactactcaaggagtaaaatctctgttaccagcatg 4380  
QY 4602 tatgtgaagaggttcctcatctccagcagtcagaatggtccatcaagttgacctctttt 4661  
|||||  
Db 4381 tatgtgaagaggttcctcatctccagcagtcagaatggtccatcaagttgacctctttt 4440  
QY 4662 cagaatgycaaaagtaaaagtttctcagggaatacagaactccttcacacctgtgtgaac 4721  
|||||  
Db 4441 cagaatgycaaaagtaaaagtttctcagggaatacagaactccttcacacctgtgtgaac 4500  
QY 4722 tctctaagaccacggttaactactgcacttgcgaattcaacccccagagttggtgcac 4781  
|||||  
Db 4501 tctctaagaccacggttaactactgcacttgcgaattcaacccccagagttggtgcac 4560  
QY 4782 cagattgcacctgagagtgaggttctggtctgcagagcagagaccttactgactcgag 4841  
|||||  
Db 4561 cagattgcacctgagagtgaggttctggtctgcagagcagagaccttactgaggtggtg 4620  
QY 4842 cgaagttc 4848  
| | | |  
Db 4621 ccaactgc 4627



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 17, 2001, 17:14:18 ; Search time 22045.3 Seconds  
(without alignments)  
5573.794 Million cell updates/sec

Title: US-09-689-430-1  
Perfect score: 7944  
Sequence: 1 tggccactccctctctgcgc.....agggggggggggggggggt 7944

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl:\*  
1: gb\_ba1:\*  
2: gb\_ba2:\*  
3: gb\_ba3:\*  
4: gb\_in1:\*  
5: gb\_in2:\*  
6: gb\_in3:\*  
7: gb\_om:\*  
8: gb\_ov:\*  
9: gb\_pat1:\*  
10: gb\_pat2:\*  
11: gb\_ph:\*  
12: gb\_pl1:\*  
13: gb\_pl2:\*  
14: gb\_pl3:\*  
15: gb\_pl4:\*  
16: em\_ba1:\*  
17: em\_ba2:\*  
18: em\_fun:\*  
19: em\_htgo\_hum:\*  
20: em\_htgo\_inv:\*  
21: em\_htgo\_rod:\*  
22: em\_htg\_hum1:\*  
23: em\_htg\_hum2:\*  
24: em\_htg\_hum3:\*  
25: em\_htg\_hum4:\*  
26: em\_htg\_hum5:\*  
27: em\_htg\_hum6:\*  
28: em\_htg\_hum7:\*  
29: em\_htg\_hum8:\*  
30: em\_htg\_inv1:\*  
31: em\_htg\_inv2:\*  
32: em\_htg\_other:\*  
33: em\_htg\_rod:\*  
34: em\_hum1:\*  
35: em\_hum2:\*  
36: em\_hum3:\*  
37: em\_hum4:\*  
38: em\_hum5:\*  
39: em\_hum6:\*  
40: em\_hum7:\*  
41: em\_in:\*  
42: em\_om:\*  
43: em\_or:\*

44: em\_ov:\*  
45: em\_pat:\*  
46: em\_ph:\*  
47: em\_pl:\*  
48: em\_ro:\*  
49: em\_sts:\*  
50: em\_sy:\*  
51: em\_un:\*  
52: em\_vi:\*  
53: gb\_sts1:\*  
54: gb\_sts2:\*  
55: gb\_sts3:\*  
56: gb\_sy:\*  
57: gb\_un:\*  
58: gb\_v11:\*  
59: gb\_v12:\*  
60: gb\_htg1:\*  
61: gb\_htg2:\*  
62: gb\_htg3:\*  
63: gb\_htg4:\*  
64: gb\_htg5:\*  
65: gb\_htg6:\*  
66: gb\_htg7:\*  
67: gb\_htg8:\*  
68: gb\_htg9:\*  
69: gb\_htg10:\*  
70: gb\_htg11:\*  
71: gb\_htg12:\*  
72: gb\_htg13:\*  
73: gb\_htg14:\*  
74: gb\_htg15:\*  
75: gb\_htg16:\*  
76: gb\_htg17:\*  
77: gb\_htg18:\*  
78: gb\_htg19:\*  
79: gb\_htg20:\*  
80: gb\_htg21:\*  
81: gb\_htg22:\*  
82: gb\_htg23:\*  
83: gb\_htg24:\*  
84: gb\_htg25:\*  
85: gb\_pr1:\*  
86: gb\_pr2:\*  
87: gb\_pr3:\*  
88: gb\_pr4:\*  
89: gb\_pr5:\*  
90: gb\_pr6:\*  
91: gb\_pr7:\*  
92: gb\_pr8:\*  
93: gb\_pr9:\*  
94: gb\_rot1:\*  
95: gb\_rot2:\*  
96: gb\_in4:\*  
97: gb\_pr10:\*  
98: em\_ba3:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4411.6	55.5	9354	9 AR003585	AR003585 Sequence
2	4357.6	54.9	4670	9 AR110040	AR110040 Sequence
3	4148.8	52.2	4278	10 I08644	I08644 Sequence 4
4	4147.2	52.2	4281	10 I08643	I08643 Sequence 3
5	4085	51.4	4548	10 I08642	I08642 Sequence 2
6	4085	51.4	4551	10 I08641	I08641 Sequence 1
7	3748	47.2	5035	9 AR034084	AR034084 Sequence
8	3748	47.2	5035	9 AR071306	AR071306 Sequence



9	3189.8	40.2	4334	5	AR029098	AR029098 Sequence
10	2830.6	35.6	6786	56	CVU89927	U89927 Cloning vec
11	2809.4	35.4	3179	56	CGEM13ZP	X65315 Cloning vec
12	2809.4	35.4	3197	56	CVPGEM3ZP	X65306 Cloning vec
13	2809.4	35.4	3221	56	CGEM11ZP	X65313 Cloning vec
14	2794.2	35.2	4118	9	A93963	A93963 Sequence 17
15	2794.2	35.2	4118	9	A93964	A93964 Sequence 18
16	2781	35.0	3582	56	SYNEBOMCS	M68846 Cloning vec
17	2781	35.0	3642	56	SYN260BLA	M29362 Plasmid pRT
18	2764.2	34.8	4076	56	AF013258	AF013258 Expressio
19	2764.2	34.8	4905	56	ASPGST	X78316 Artificial
20	2762.6	34.8	6094	9	A64973	A64973 Sequence 3
21	2758.8	34.7	3984	10	I15353	I15353 Sequence 1
22	2741.8	34.5	5676	56	CVPSELECT	X63334 Cloning vec
23	2719	34.2	5858	56	CVU47102	U47102 Cloning vec
24	2699.2	34.0	6101	56	CVPSIT	U11309 Phagemid cl
25	2637.8	33.2	6162	10	E14304	U14304 pUC118. 7/1
26	2637.8	33.2	8532	56	U02452	U02452 Cloning vec
27	2637.8	33.2	9593	10	E16678	E16678 All sequence
28	2636.2	33.2	4696	9	AR069374	AR069374 Sequence
29	2636.2	33.2	5158	9	AR069375	AR069375 Sequence
30	2636.2	33.2	7607	10	I14516	I14516 Sequence 19
31	2636.2	33.2	19307	9	AR093392	AR093392 Sequence
32	2631.6	33.1	5461	56	U02450	U02450 Cloning vec
33	2631.6	33.1	5538	56	U02446	U02446 Cloning vec
34	2631.6	33.1	5538	56	XXU02440	U02440 Cloning vec
35	2631.6	33.1	11958	9	AX057269	AX057269 Sequence
36	2631.4	33.1	9208	56	AF064067	AF064067 Expressio
37	2631.4	33.1	9211	56	AF064066	AF064066 Expressio
38	2631.4	33.1	9215	56	AF049063	AF049063 Expressio
39	2631.4	33.1	9956	56	AF049064	AF049064 Expressio
40	2625.8	33.1	4054	9	A81878	A81878 Sequence 9
41	2621	33.0	3632	56	CVU47121	U47121 Cloning vec
42	2621	33.0	4006	56	CVU47119	U47119 Cloning vec
43	2605.8	32.8	4910	56	ASAJ156	AJ000156 Artificial
44	2605.8	32.8	4951	56	ASAJ154	AJ000154 Artificial
45	2605.8	32.8	5284	56	ASOURESN	Y11035 Artificial

## ALIGNMENTS

RESULT	1				
LOCUS	AR003585				
DEFINITION	AR003585	9354 bp	DNA	PAT	04-DEC-1998
ACCESSION	Sequence 2 from patent US 5744326.				
VERSION	AR003585				
KEYWORDS	AR003585.1	GI:3964844			
SOURCE	.				
ORGANISM	Unknown.				
	Unknown.				
	Unclassified.				
REFERENCE	1 (bases 1 to 9354)				
AUTHORS	Ill,C.R. and Blottingmaler,S.				
TITLE	Use of viral CIS-acting post-transcriptional regulatory sequences to increase expression of intronless genes containing near-consensus splice sites				
JOURNAL	Patent: US 5744326-A 2 28-Apr-1998;				
FEATURES	Location/Qualifiers				
source	1..9354				
	/organism="unknown"				
BASE COUNT	2506 a 2239 c 2161 g 2448 t				
ORIGIN					

	Query Match	55.5%	Score 4411.6	DB 9	Length 9334
	Best Local Similarity	99.7%	Pred. No. 0		
	Matches 4420	Conservative	0	Mismatches 14	Indels 0
				Gaps 0	
QY	415	ccacatgcgaatagacgtccacactggtctctctctgacctttgacatcctggtta	474		
D8	2960	ccacatggtttatgacgtccacactgcttttctgacctttggacattctgcgttga	3019		

QY	475	gtgcacccaagaagtactactctggtgtcagttggaacgtcaatggagctctatgacaag	534
Db	3020	gtgccaccagaagaatctactctgggtgcagtggaacgtatgatggacattatgacaaagtg	3078
QY	535	alctcgtgtgaagctgcctctgtgaagccaagatctccctccatagatgccaacatcttccat	594
Db	3080	attctcgtgtgacctgcctctgttgagcccaagatttctctccttaagtgccaataattctttccat	3138
QY	595	tcaacacctcagtcgtgtacaaaaagacctcgtttttagaatcaatcagtgctcaacttcca	654
Db	3140	tcaaacacctctcagtcgtgtacaaaaagacctcgtttttagaatcaatcagtgctcaacttcca	3198
QY	655	acatcgtccaagccaagggccaccctggatgggtctcgaagtcctaccatccagcgctcgag	714
Db	3200	acatcgtccaagccaagggccaccctggatgggtctcgaagtcctaccatccagcgctcgag	3258
QY	715	tttatgatacagtggtcatctacacttaagacaatggtctccactcctgtccactgtcatcg	774
Db	3260	ttttatgatgacgtgctgtctattacattgaagacatggcttcccatcctgttcagctttcag	3318
QY	775	ctgtgtgtgtatccctactgynaagcctctctgaaggagctgaatagtatgaacagaccgctc	834
Db	3320	ctgtgtgtgtatccctactgynaagcctctctgaaggagctgaatagtatgaacagaccgctc	3378
QY	835	aaagggagaaaagaatgataaagctctctccctgtgtgaagcaatactgtctggcag	894
Db	3380	aaagggagaaaagaatgataaagctctctccctgtgtgaagcaatactgtgtccagcag	3438
QY	895	tcctbaaagagaatggtgtccaaatggccctctaccacacgtgccttactactaatcctt	954
Db	3440	tcctbaaagagaatggtgtccaaatggccctctgaaccactgtgccttactactaatcctt	3498
QY	955	ctcatgtgagacctgtgtlaaaaagctctgaatctcaagcctcatctgagccctactaglatgta	1014
Db	3500	ctcatgtgagacctgtgtlaaaaagctctgaatctcaagcctcatctgagccctactaglatgta	3558
QY	1015	gagaaagagatctctggccaaggaagaaacacagacctgtgccaataattactacttltg	1074
Db	3560	gagaaagagatctctggccaaggaagaaacacagacctgtgccaataattatctacttltg	3618
QY	1075	ctgtatttgaatgaaagggaagaaagttgtgacacccaagaaacactcctgtatgacagata	1134
Db	3620	ctgtatttgaatgaaagggaagaaagttgtgacacccaagaaacactcctgtatgacagata	3678
QY	1135	ggagatgctgcatctcgtcctgggctctgagcctlaaaatgacacagatgcaatgtgtatgaaca	1194
Db	3680	ggagatgctgcatctcgtcctgggctctgagcctlaaaatgacacagatgcaatgtgtatgaaca	3738
QY	1195	ggtctctgcacagctctgtatgtgatgccaagaaatcgtatctatgtgcatgtgatgtgaa	1254
Db	3740	ggtctctgcacagctctgtatgtgatgccaagaaatcgtatctatgtgcatgtgatgtgaa	3798
QY	1255	tgggacacacacctccgaagtgcacaaatctcctcgaaggtcaacaaatctctgtgaga	1314
Db	3800	tgggacacacacctccgaaggtgcacaaatctcctcgaaggtcaacaaatctctgtgaga	3858
QY	1315	acacatcgcccaaggtctcctctgaaatctcgccaataaacttccctactgtccaacactc	1374
Db	3860	acacatcgcccaaggtctcctctgaaatctcgccaataaacttccctactgtccaacactc	3918
QY	1375	tgatgtgacctgtgacagttctctactgttttgtcatactctctccccaacaatgtatgaca	1434
Db	3920	tgatgtgacctgtgtgacagttctctactgttttgtcatactctctccccaacaatgtatgaca	3978
QY	1435	tgggaagcttatgtcaaagtgtgacagctgtctcagaagaaacccaactcaagaaatgaanaata	1494
Db	3980	tgggaagcttatgtcaaagtgtgacagctgtctcagaagaaacccaactcaagaaatgaanaata	4038
QY	1495	atgaagaagcgggaagactatgaatgatatcttactgtactcgtgaatgtgtgtgtcagct	1554
Db	4040	atgaagaagcgggaagactatgaatgatatcttactgtactcgtgaatgtgtgtgtcagct	4098
QY	1555	tgtatgtatgaaacactctcctcctcttataccaataatcgctcgaagttgtgccaagaagcatctca	1614



|||||  
Db 4100 TTGATGATGACAACTCTCTCTTATTCOAATTCGCTCACTGTCGAAGAAGATCCGA 4159  
OY 1615 aaacttggtaacttaacttctgtctgaagaggaacttggacacatbctcccttaagtc 1674  
Db 4160 AAACCTTGGGTACATTACCTTGGCTGGAAGAGAGACTGGACATCATCTCCCTTAGTCC 4219  
OY 1675 tgcgccccgatcacagaagttataaagttcaattttgaacaaatggccctcagagatg 1734  
Db 4220 TCGCCCCGATACAGAAAGTTTAAAGTCATATTGACAAATGGCCCTCAGCGGATTTG 4279  
OY 1735 gtaaggaatlacaaaaaagttccgaattatgycacacagatgaaaccttlaagactcg 1794  
Db 4280 GTAGGAAGTACAAAAGATCCGATTTATGACATACAGATGAACCTTTAAGACTCGTG 4339  
OY 1795 aagcattcagatgaatcacgaagactcttggagaccttacttaatggggaagtltgagaca 1854  
Db 4340 AAGCTATTCAGATGAATCAGGAATCTTGGACCTTTACTTTATGGGAAGTTGGAGACA 4399  
OY 1855 caatgtatataataatlaagaatcaagaagcagacacataataacttaacctcagaa 1914  
Db 4400 CACTGTGATTTATTTAAGATTCAGACAGACAGACATATACATCTACCTCAGGAA 4459  
OY 1915 tcaatgatgtccgtctctgtatltcaagagatltacaaaaglytlaaacaatttgaag 1974  
Db 4460 TCACGTAGTCCGCTCTTGTATTCAGAGATACCAAAAAGTGTAAACATTTGAAGG 4519  
OY 1975 atttccaattctgcagagagaataatcaataataatgacagtgactgtatagaagatg 2034  
Db 4520 ATTTTCCAAATTCGACAGAGAAATTTCAATATAATGACATGTAGTGAAGATG 4579  
OY 2035 ggcacaactaatcaatcctcgtgtccgcagccgctatctactagtcttlaatag 2094  
Db 4580 GCGCAACTAATACAGATCCTCGGTGCTGACCCGCTATTACTCTAGTTTCGTAATAGG 4639  
OY 2095 agaagatcagctcaggaactatlggcccctccatcctcgtctacacaaagaatctgag 2154  
Db 4640 AGAGATCTGATCTCAGAGACTCATTTGGCCCTCTCATCTGCTACAAAGATCTGAG 4699  
OY 2155 atcaaaagggaacacagtaattgtcagacaagagaatgtatcctgttctctgtatttg 2214  
Db 4700 ATCAAGAGGAAACCCAGATTAATGTGACAGAGGAATGTATCTCTGTTTCTGTATTG 4759  
OY 2215 atgaagaacgaagctgtgacctcacagagaatatacaagcttctcccaatccagctg 2274  
Db 4760 ATGAACAACGGAAGCTGTGATCTCAAGAGATATACACGCTTTTCCCAATCCAGTGG 4819  
OY 2275 gaatgcagcttgaagatccagaagttccaagcctccaacatcatgacagca tcaatggct 2334  
Db 4820 GAGTGCACCTTGAGGATCCAGATTCACAAGCCTCCAMCATCATGCACAGCATCATGCT 4879  
OY 2335 atgttttgatagtttgaagtgttgaagtttggatgagatgtgacacatgtaacatc 2394  
Db 4880 ATGTTTGTGATAGTTTGCAAGTTGTCAAGTTTGTTCATAGAGTGGCAATCTGCATATTC 4939  
OY 2395 taagcatggagcacaagctactccttctcgtctctcctcctcgtgatatccctcaac 2454  
Db 4940 TAAAGATTGGAGACAGACTGACTCTTCTGTCTTCTTCTGTGATATACCTTCAAC 4999  
OY 2455 acaaaatggtctatgaagacaacactacacctatccca tctcagagagaatcgtctca 2514  
Db 5000 ACAAAATGCTATGAAACACACACTCACCTATTCCTCATGAGAGAAATGCTCTCA 5059  
OY 2515 tgtcgaatgaanaacagttatagatctcgggtgtgcgaacactcagacttccggaaca 2574  
Db 5060 TGTGATGAGAAACCCAGCTCTATGATTTGTGGGTGCAACACTCAACATTTTGGGAACA 5119  
OY 2575 gaagcatgacccgtctactgaagttctctagttgtgacaagaacactgttgaattacag 2634  
Db 5120 GAGGATGACCGCTTACTGAAAGTTTCTAGTTGTGACAAACACTGCTGATATTATTCG 5179  
OY 2635 aggaacgttatgaagatatctcagcatactgtcgtgagtaaaaaaatgcaatgcaacca 2694  
|||||

Db 5180 AGACATTTATGACATATTTACAGCATCTTGTGAGTAAAAACAATGCATTTGAACCA 5239  
OY 2695 gaagcttctccagaatltcaagaacccctagcaactaggcaaaagcaatttaatgccacc 2754  
Db 5240 GAAGCTTCTCCAGAAATTCAGACACCCCTAGCACAGCAAAAGCAAAATTAATGCAACC 5299  
OY 2755 caccagcttgaagaacgcatcaacgggaataaactcgtactactccttcaatcaatgaag 2814  
Db 5300 CACCGATTTGAACGCCATACACGGGAATTAATCTGACTACTCTTCTGTCAGATCAG 5359  
OY 2815 aggaattgactatgatalgataccatalcagttgaatltgaagaagaagaatttgaacatt 2874  
Db 5360 AGGAATTTGACTATGATATACATATCAGTTGAATGAAGAAGAGATTTTGACATTT 5419  
OY 2875 atgatgaagatgaanaatcaagagcccccgaagcttccaagaagaaacagacacatttta 2934  
Db 5420 ATGATGAGGATGAANAATGAGAGCCCCGACCTTTCAAAAACACAGACATATTTTA 5479  
OY 2935 ttgctgcagttgagaggtctcgggata tggatgagtagctccccaatgttctaagaa 2994  
Db 5480 TTGCTGCAGTGAGAGGCTCTGGATTTATGGATGAGTAGCTCCCAATGTTCTAAGAA 5539  
OY 2995 acaaggctcagaatgagcagttccctcagttcaagaagttgtlttccaagaaattactg 3054  
Db 5540 ACAGGGCTCAGAGTGGCAGTGTCCCTCAGTTCAAGAAAGTTGTTTCCAGGAATTTACTG 5599  
OY 3055 atggtccttactcagcccttataccgttgaagaactaaatgaaacatttggactccgg 3114  
Db 5600 ATGGCTCTTTACTCAGCCCTTATACCGTGAAGAACTAAATGAACATTTGGACCTCGG 5659  
OY 3115 ggcacataatagacagaagttgaagataatcatatgtaacttccaagaatcagagct 3174  
Db 5660 GGCATATATTAAGACAGAAAGTTGAAGATATATCATGTGTAATTTCAAAATCAGGCT 5719  
OY 3175 ctgblccctattccttctactttagccttattcttlaatgaagaatcagaagcagaag 3234  
Db 5720 CTCGTCCTATTCTCTTATCTACCTTATTTCTTATGAGGAAGATCAGAGCAAGAG 5779  
OY 3235 cagaacctagaanaaacttltcaagcctlaatgaacccaacttacttltgggaagtgc 3294  
Db 5780 CAGAACCTTGAAGAAACTTGTGTCAAGCCTAATGAACCAAACTTACTTTGGAAAGTC 5839  
OY 3295 aacatcataatgcaacccactaaagaatgagttgactcgtcaagcctcgtgctattctctg 3354  
Db 5840 AACATCATATGCGACCCCACTAAAGATGATTGACTGCAAAAGCCTGAGCTTATTTCTG 5899  
OY 3355 atgtgacctggaanaaagatgtgactcagcagcctgatgtgagcccttctgctgcaca 3414  
Db 5900 ATGTTGACCTGGAAGAAAGATGTGACTCAGGCTGATGTGACCCCTCTGTGCTGCCACA 5959  
OY 3415 cttaacaacttgaacccctgtcatgtggagacaagtgaaagtacaagaatttgcctgtttt 3474  
Db 5960 CTAACACACTGAACCCCTGCTCATGTGAGACAAAGTACAGTACAGAAATTTGCTCTGTTT 6019  
OY 3475 tcaccatctctgatatgagccaaagctgttactcactcgtgaataatltgaagaagaactgca 3534  
Db 6020 TCACACTTTTGTATGAGAGCAAAAGCTGTCTTCTACAGTGAATAATGGAAGAAACTGCA 6079  
OY 3535 gggctccctgcaatataccagaatgaagatccactttaaagagaattatcgtctcag 3594  
Db 6080 GGGCTCCCTGCAATATCCAGATGGAAGATCCCACTTTTAAAGAAATATGCTTCCATG 6139  
OY 3595 caatcaatggtctataatagatatcacactcgtcttagtaatgttccagaatcaagaatga 3654  
Db 6140 CAATCAATAGGCTATATATGATATACACTGCTGCTTATGTAATGCTCAGGATCAAGGA 6199  
OY 3655 ttcgatgttactcgtcagcaatggcagcaatgaaacatccatctatltactcag 3714  
Db 6200 TTGATGATATCTCTCAGCATGAGCAATGAGCAATGAAACATTCATTTATTTCACTG 6259  
OY 3715 gacatgttctcactgtacgnaaaaaagagagatlaaaatgagcgtgtacaaatctctatc 3774  
Db 6260 GACATGTTGTTCACTGTACGAAAAAAGAGAGTAAATGGCACTGTACAACTCTATTC 6319  
|||||



QY 3775 caggtctttttagacagtggaatgtacacacccaaagcttggaatttggcgggtggaat 3834  
|||||  
Db 6320 CAGGCTTTTACAGACAGTGAATGTTACATCCAAAGCTGGAATTTGGGGGTGGAAT 6379  
QY 3835 gacctatctggcagacatcacatgctgggaatgagcacactttcttggtgtagacaata 3894  
|||||  
Db 6380 GCTTATTGGGACGACATCATGCTGGGATGAGCACACTTTTTCGGTGTACAGCAATA 6439  
QY 3895 agtgcacacacccctgggaatggctctggaacacatgagaatttcagattacagctt 3954  
|||||  
Db 6440 AGTGCACACCTCCCTGGGAATGGCTTCTGACACATTAGAGATTTTCAGATTACAGCTT 6499  
QY 3955 caggacaaataggaagatgggcccacaaagctggcagacatcatalcgcgaataca 4014  
|||||  
Db 6500 CAGGACAAATATGACAGTGGGCCCAAGCTGGCAGACTTATTTCCGGATCAATCA 6559  
QY 4015 atgcttggaagcaaggaagcccttctctggaatcaaggatgagctggttggaacaa 4074  
|||||  
Db 6560 ATGCTTGAGGACCAAGAGCCCTTTTCTTGATCAAGGTGATCTGTGGCACCAATGA 6619  
QY 4075 ttattcaggagcaatgaagaccagggtggcccgfcaagaagttccacacccctcatalcttc 4134  
|||||  
Db 6620 TTATTACAGGCAATCAGACCCAGGGTCCCGTACAGAGTTCTCCAGCCTCTACATCTCTC 6679  
QY 4135 agttatcatcatgatatagctcttgaatgggaagagtgagacatcagagaatccca 4194  
|||||  
Db 6680 AGTTTATCATCATGTATAGTCTTGATGGGAGAGAGTGGCAGACTTATTCAGGAAATTTCA 6739  
QY 4195 ctggaagcttaagtgctctcttggcaatgagatcatcctgggataaacaacataatt 4254  
|||||  
Db 6740 CTGGAACCTTATATGCTCTTCTTGGCAATGTGATTCATCTGGGATTAACCAATATTT 6799  
QY 4255 ttaacccctcaatcttgctgacatcacccgttggacccaactcttataagattcgca 4314  
|||||  
Db 6800 TTAAACCTCCCAATTATTTGCTCGATACATCCGTTTGCACCCCACTCTTATAGCATTTGCA 6859  
QY 4315 gcaactctcgacatggagtgtgagctgtgatatataatagttgcaagcatgcatgtgga 4374  
|||||  
Db 6860 GCACCTTTCGCAATGAGTGTGATGGGCTGTGATTTAAATAGTTGACACATGCCATTGGAA 6919  
QY 4375 tggagagtaagaacatcacatgagatgcagataactgctcattccatcacataatt 4434  
|||||  
Db 6920 TGGAGAGTAAGCAATATCATGATGACACAGATTACTGCTTATCTTACCAATATGT 6979  
QY 4435 ttgcaactgtgtctccttcaaaaagctcgacatcgaactcgaaggaagatgagcttga 4494  
|||||  
Db 6980 TTGCCACCTGTGCTCTTCAAAAAGCTCGACTTCACCTCCAAAGGAGAGTAATGCTTGA 7039  
QY 4495 gaactcagtgaaataatccaaaagtggtgcaagtggaactccagaagaatgaaag 4554  
|||||  
Db 7040 GACCTCAGGTGAATATCCAAAAGAGTGGCTGCAAGTGAAGTCCAGAGACAAATGAAG 7099  
QY 4555 tcaaaagtgtaactacacaggaagtaaaaatctctgtcttaccagacatgatatggaagat 4614  
|||||  
Db 7100 TCACAGGAGTAACCTACTCAGGAGTAATAATCTCTGCTTACAGCAATGTATGGAAGAGAT 7159  
QY 4615 tccatcatcagcagcagtaagaatggacatcaatggaactccttcttcagaatggaag 4674  
|||||  
Db 7160 TCCATCATCTCCAGCAGTCAAGATGGCCATCATGTGACTCTTTTTCAGAATGCAAG 7219  
QY 4675 taaagtttctcaaggaaatcaagaactcctcacaacctggtggaatctctagaaccac 4734  
|||||  
Db 7220 TAAAGGTTTTCAGGGAATCAAGACTCTTACACACTGTGGTAAGCTCTTAGAACCCAC 7279  
QY 4735 cgttactgactgctactcttgaaatcaacccccagaagttgggtgcaacgaattgccccta 4794  
|||||  
Db 7280 CGTTACTACTCTGCTTACCTTCAATTCACCCAGAGTTGGGTGCAACAGATTGGCCCTGA 7339  
QY 4795 ggaatggaagttctgggtctgagagcacaagaactctactgactgagcgagtttc 4848  
|||||  
Db 7340 GGATGGAGGTTCTGGGCTGCGAGGACAGACCTCTACTGAGGGTGGCCACTGCG 7393

RESULT 2  
LOCUS AR110040 4670 bp DNA PAT 14-FEB-2001  
DEFINITION Sequence 41 from patent US 6114148.  
ACCESSION AR110040  
VERSION AR110040.1 GI:12826316  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source  
BASE COUNT 1325 a 1072 c 1036 g 1237 t  
ORIGIN  
Query Match 54.9%; Score 4357.6; DB 9; Length 4670;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 4363; Conservative 0; Mismatches 9; Indels 0; Gaps 0;  
QY 477 gccaccagaagatactacctgggtgagtgagtgagactgcatggaactatgcaagtgtat 536  
|||||  
Db 84 GCCACGAGAAGTACTACTGCTGGTGCAGTGAACGTGTATGAGACTATATGCAAAAGTAT 143  
QY 537 ctctggtgacgtgcttggtggaacgaattctcctctagatgagtgcaaaatctttccattc 596  
|||||  
Db 144 CTCGGGAGCTGCTGCTGAGCAGCAAGATTCTCCCTAAGTGCCAAAATCTTTTCATTC 203  
QY 597 aacacctcagtggtgacaaaagaagctcgtttgtagaatcaaggttcacatttcaac 656  
|||||  
Db 204 AACACCTCAGTGTGTCACAAAAGACTGTTGTGTAATTCACGATACCTTTTCAAC 263  
QY 657 atcgcctcaagcgaagccacccctggaatgggtctgctgaagtcctcaccatcgaagctgag 716  
|||||  
Db 264 ATCGCTAACCAAGGCAACCTGATGGGTGCTGCTAGTGCCTACCATCAGGCTGAGGTT 323  
QY 717 tatgatacagtgatcaacttaagaacatggtctcccatcctgtcaagttcaatgct 776  
|||||  
Db 324 TATGATACAGTGTGATTCACCTTAAGAACATGAGCTTCCATCTGTCAGTCTTCATGCT 383  
QY 777 gttgtgtatcctactcaggaagctctcgaagagcttgaatgatgatacagaacagttca 836  
|||||  
Db 384 GTTGGGTATCTCTACTGGAAAGCTTCTGAGGAGCTGATATGATGATCAGCCAGTCA 443  
QY 837 agggagaagaagatgataaagtcttccctggctggaagccacatacatatgtctgcaagctc 896  
|||||  
Db 444 AGGGAGAAAGAGATGATTAAGTCTTCCCTGTGGAAGCCATACATATGTCTGGAGGTC 503  
QY 897 ctgaaagaagatggttccaatggtcctcgaaccatggtcctactactacatattctc 956  
|||||  
Db 504 CTGAAGAAGAAAGTGCATATGCTGCTGACCCACATGCTGCTTACCTACTCATATCTTCT 563  
QY 957 catgtgaccttggttaaaagacttgaactcaggccctcatgtggaagccctactgattgaga 1016  
|||||  
Db 564 CATGTGAGCTGTGTAAMAACTTGAAATTCAGGCTCATATTGGAGCCCTACTGTATGTATGA 623  
QY 1017 gaaagagcttggaacgaagaaagacacagacactgtgacaaattatatactacttctgct 1076  
|||||  
Db 624 GAAAGGAGTCTGGCCAGGAAAGAACACAGACACTTGCACAAATTTATATCTTTTCT 683  
QY 1077 gtaattgatgaaggaagatgtggaactcagaacaaagaactccttgaatgcagatagg 1136  
|||||  
Db 684 GTATTGTAGTAGAGGAAAGATTGGCAGCTCAGAAACAAAGAACTCTTGTATGACAGATAGG 743  
QY 1137 gatgtcgaatcgtgtccggcctggtccttaaatgacacaaatgataatgataatgataatg 1196  
|||||  
Db 744 GATGCTGATCTGCTCGGCTGGCTTAAATGACACAGTCAATGATTAATGTATTAACAGG 803



QY	1197	lclctcgcaagtcctgaltggtatggtccacaggaatlcagctcattgagctgaltggaatg	1256
Db	804	TTCTTCCCGAGGCTCTGATTTGGATGGCCACAGGAATCAGTCATTTGGCATGATTTGGAATG	863
QY	1257	ggcacaccctccgaagtgcacccaatatcttcctcgaaaggtccacacattctctgtggaagac	1316
Db	864	GGCACACCTCCGAATGTGACCTAAATATTCCTGGAAGGTACACATTTCTTTGGAGAAC	923
QY	1317	catcgccagagctccctctggaatctcgcaccaatacttccctactctgctccaacactctg	1376
Db	924	CATTCGCCAGGCTCCTTTGGAAATCTCGCAATTACTTCTTCTTCTCTCAAAACACTTGG	983
QY	1377	atggaccttggaagtlcttactgtttgtcatalctcttcccacaacatgatlgaacatg	1436
Db	984	ATGGACCTTGGACAGTTTCTACTGTTTGTTCATATCTCTTCCACCAACATGATGGCATG	1043
QY	1437	gaagcttatgtaaagttaacagctgttccagaggaaccccaactaagaatgaaaaataat	1496
Db	1044	GAAGCTTATGTCAAAAGTAAACAGCTGTCCAGAGAACCCCACTACGAAGAAATTAAT	1103
QY	1497	gaagaagcggaagactatgatgatgtactactgtatctcgaaatgtgctgagctt	1556
Db	1104	GAAAGAGCGGAAGACTATGATGATCTTACTGATTTCTGAAATGATGTGGTGCAGTTT	1163
QY	1557	gatgatgcacctctccctccttalcacaatctgcctcagctgtgccaagaagcaltccaa	1616
Db	1164	GATGATGACAACTCTCCCTTCTTTATCCAAATTCGCTCAGTTGCCAAGAAGCTCTAA	1223
QY	1617	acttgggtacatataactgtctgtcgaagagggagactggactatgctcccttaagcttc	1676
Db	1224	ACTTTGGGTACATTTACATTTGCTCTGTAAGAGAGGACTGGAGCTATCTCCCTTAATGCTC	1283
QY	1677	ggccccgatgcagaagttataaazgtcaatatlttgaacaatgcccccaagcgaattgtg	1736
Db	1284	GGCCCCGATGACAGAAGTATTATAAAGTCAATTTTGAACAATGGCCCCCTCAGCGAATTGG	1343
QY	1737	aggaagtacaaaaaagtcggacttaltgycatacaacagatgaaacotttaagactgtgaa	1796
Db	1344	AGGAAGTACAAAAAAGTCCGATTTATGGCATACAGATGAACCTTTTAAGACTCGTGAA	1403
QY	1797	gctattccagcatgaatcgaagactctgtggaccttactcttataggggaagtltgagaacaa	1856
Db	1404	GCTATTTCAGCATGAATCAGGAATCTTTGGAGCCTTTACTTTATGGGGAAGTTGGAGACAC	1463
QY	1857	ctgttgatatalttaagaatcaagcaagcagacacatalaactcaccctccaggaatc	1916
Db	1464	CTGTGATATATTTATAAGATCAAGAACGACCATATTAACATCTAACCTCAGCGAATTC	1523
QY	1917	actgatgtcgtcccttctgtatctcaaggagatataccaaaaggtgtaaacatttgaagat	1976
Db	1524	ACTGATGTCCGCTCTTTGATTTCAAGGAGATTAACCAAAAGGTGTAAACATTTGAAGGAT	1583
QY	1977	tttccaactctgcagaggaataattccaataataatggaagtgagcttagaagaatgag	2036
Db	1584	TTTTCCATTTCTCCGAGGAGAAATATTCAAAATTAATGACACGTGCTAGAAAGTGGC	1643
QY	2037	ccaactaaatcagatcctcgtgctgcagccgctaltactagttctgltlaatatgagag	2096
Db	1644	CCAACGTAATTCAGATCTCGTGGTGGCCGAGCCCGCTATTACTAGTTTGGTTAATATGAG	1703
QY	2097	agaagcttagactcgaagatcaltgtgccccctccctcaltctgctcaaaaaaactgtagat	2156
Db	1704	AGAGATCTTAGCTTCAGGACTCATTTGGCCCTCTCTCATCTGTCTACAAAAGATTTGTAGAT	1763
QY	2157	caaagaagaaacacataatgtcagacaagaggaatgtcatccgtttctgtatctgag	2216
Db	1764	CAAGAAGGAAACACAGATATATGTTCAGACAAAGAGGAATGTCACTCTGTTTCTGTGATTTGAT	1823
QY	2217	gagaaacgaagctgtgtacctcacaagaatatatacagctctctccccaatccaagctgaa	2276
Db	1824	GAGAACCGGAAGCTGTACTCTCAAGAGAAATATAAAGCTTTCTCCCAATCCAGCTGGA	1883

QY	2277	gtcagccttgagatccaaggtccaaagctccaacataigcaagcaatcaatgctat	2338
Db	1884	gtgcagcttgaggatccaaggtccaagctccaacatcatcacagcatcatggtat	1943
QY	2337	gttttgaatgttcaggtgttcaagttgtgttcagaaggtggcaactgtaacta	2396
Db	1944	gtttttgattactttgcagttgtcagtttgttttcattgagagtgccatctgta	2003
QY	2397	agcalttgagacagactgtccttcctgtccttccttccttgatatacctccaac	2456
Db	2004	agcatwtgagacagacagctgacttccttctgtctctctctgattatcttcaaac	206
QY	2457	aaatggtctatgaagacaactcaactatcccatctcagagaactgtctcatg	2516
Db	2064	aaatggtctatgaagacaactcaactatcccatctcagagaactgtctcatg	2122
QY	2517	tcgagtgaataacccaagttcatatgattctcgggtgcacaactcagacttggaga	2576
Db	2134	tcgagtgaataacccaagttcatatgattctcgggtgcacaactcagacttggaga	2183
QY	2577	ggcaagaccgcttaccgaaggtttctagttctgcaagaacaactgtaattacag	2633
Db	2184	ggcattgacccgcttactgtaaggtttctagttgttgacaaagacacttgattat	2243
QY	2637	gacagttatgaagatattccaactctgtctgtagtaaaacaatgccaatgaacca	2696
Db	2244	gacagttatgaagatattccaactctgtctgtagtaaaacaatgccaatgaacca	2303
QY	2697	agctcttcccgaattccaagaacacctagcaactaggcaagaatattaatgcacca	2755
Db	2304	agctcttcccgaattccaagaacacctagcaactaggcaagaatattatccaccca	2363
QY	2757	ccagctctgaaagccaacaaggaaataacccgttaactcttcagtcagatcaag	2818
Db	2364	ccagctctgaaagccaacaaggaaataacccgttaactcttcagtcagatcaag	2423
QY	2817	gaatltgaactatgaatgataccatacatgattgaatgaagaagaagatttgacat	2876
Db	2424	gaatltgaactatgaatgataccatacatgattgaatgaagaagaagatttgacat	2483
QY	2877	gatgagatgaataatcagagcccccagagcttccaagaacaaccgaacatttat	2933
Db	2484	gatgagatgaataatcagagcccccagagcttccaagaacaaccgaacatttat	2543
QY	2937	gctcagatggagagagccttggaattatgagatagtagtcccccacaatgttcaaga	2996
Db	2544	gctcagatggagagagccttggaattatgagatagtagtcccccacaatgttcaaga	2603
QY	2997	agggctcagagtggcagtgctccctcagatccaagaagaattgttttccagaattac	3056
Db	2604	agggctcagagtggcagtgctccctcagatccaagaagaattgttttccagaattac	2663
QY	3057	ggctccttactcagcccttataccgttgaggaactaaatgaacatttggactcctgg	3116
Db	2664	ggctccttactcagcccttataccgttgaggaactaaatgaacatttggactcctgg	2723
QY	3117	ccatataaagagaagaagttgaagaataatacatgtaacttccaagaatcagccct	3176
Db	2724	ccatataaagagaagaagttgaagaataatacatgtaacttccaagaatcagccct	2783
QY	3177	cgctccattctcctcatattcagaactatttcttatgagaagatcacagagcaagaca	3236
Db	2784	cgctccattctcctcatattcagaactatttcttatgagaagatcacagagcaagaca	2843
QY	3237	gaaccttgaaaaaactttgtcaagcctaatgaaaccaaacttaactttggaagtcaa	3296
Db	2844	gaaccttgaaaaaactttgtcaagcctaatgaaaccaaacttaactttggaagtcaa	2903
QY	3297	catcatatgacccaactaaagaatgagttgactgcaaaagccttgagcttatctctg	3356
Db	2904	catcatatgacccaactaaagaatgagttgactgcaaaagccttgagcttatctctg	2963
QY	3357	gttgaactggaaaaagaatgtgcaactaagcctgattggaacccctctgctcgaact	3416



Db 2964 GTTGGACCTGGAAAAAGATGTGCACTCAGGCTGATTGGACCCCTTGTGCTGCCACACT 3023  
 Qy 3417 aacacactgaaacctgctcatalggaagacaaatgacagtaacagaaatctgctgttttc 3476  
 Db 3024 AACACACTGAACCCCTGCTCATGTGGAGACAAAGTACAGTACAGAAATTTGCTCTGTTTTTC 3083  
 Qy 3477 accacctcttgatgagaccaaaagctggtactcaactgaaataatggaagaaactgcaag 3536  
 Db 3084 ACCATCTTTGATGAGAACCAAAAGCTGTACTTCACTGAAATATGAAAGAAATCTCAGG 3143  
 Qy 3537 gctccctgcaataaccagatggaagatcccaactttaaagagaatatacgcttcatagca 3596  
 Db 3144 GCTCCCTCAATATCCAGATGGAAGATCCACTTTTAAAGAAATATGCTTCCATGCA 3203  
 Qy 3597 atcaatgctacataatgatactactaccctggtcttaagtaatgagcagaagataaagatt 3656  
 Db 3204 ATCAATGGCTACATATATGATGATCACTACTGCTTATGATGCTTACAGATCAAAAGGATT 3263  
 Qy 3657 cgatgtaactgctcagcagctggcagcaatgaaacatacattctatctcaatgca 3716  
 Db 3264 CGATGCTATCTGCTCAGCATGGGAGCAATGAACATCATTTCTATTCATTTCACTGGA 3323  
 Qy 3717 catgtgtcaactgtacgaaaaaagagagataaaatggcactgtacaaatctctacca 3776  
 Db 3324 CATGTGTTCACTGTACGAAAAAAGAGAGATTAATAATGGCACTGTACATCTCTATCCA 3383  
 Qy 3777 ggtgttttgagacagtggaatgtttaacatacacaagctgggaatttgagggtggaatgc 3836  
 Db 3384 GGTGTTTTTGACACAGTGGAAATGTTACCATCCAAAGCTGGAAATTTGGGGGTGGAAATGC 3443  
 Qy 3837 ctatctggcagacatctacatgctggaatgacacatttctctggtgtaacagcaataag 3896  
 Db 3444 CTATTTGGCGAGCATCTCATCTGAGATGAGCACACTTTTCTGTGTATACAGCAATAG 3503  
 Qy 3897 tgtcaactccctggtggaatggtctctggaacatataagaaatlttcaagtttaagcttca 3956  
 Db 3504 TGTCAGACTCCCTGTGGAAATGGCTTGTGACACATTAAGAAATTTTACAGCTTCA 3563  
 Qy 3957 ggaacatctgacagtgaggcccccaagctgagcagacatctatctcgatcaacaat 4016  
 Db 3564 GGACAAATATGACAGATGGGCCCCCAAGCTGGCAGACATTCATTATTTCCGATCAATCAAT 3623  
 Qy 4017 gctctggagcaccgaagaccttctcttgatcaagtgagatctgttggcaccgaatgatt 4076  
 Db 3624 GCTTGAGACACCAAGAGGCCCTTTTCTTGATCAAGTGTGATCTGTGGCACCAATGATT 3683  
 Qy 4077 attcaagcatalcaagaaccgaaggtgcctgcaagaatttccaagctctcaactctcaag 4136  
 Db 3684 ATTCACGGCATCAACACCAAGGTCGCCGTCAAGATTCCTCAGCCTCTACATCTCTCAG 3743  
 Qy 4137 tttaacatcatgatagtctcttgatggaagaagtgacagactatcgaagaatctcaact 4196  
 Db 3744 TTTATCATCATGATATAGTCTTATGGAAGAGAGTGGCAGACTTATTCAGGAATTTCCACT 3803  
 Qy 4197 ggaaccttaatggtctctcttgacaaatgta ttcaatctggaataaaacacataatctt 4256  
 Db 3804 GGAACCTTAATGATGCTCTTTTGGCAATGGAATCATCTGGATTAATAACATATATTTT 3863  
 Qy 4257 aaacctccaatatactgctcgatataatcgcttgcacccaactcaatataatgcatctgagc 4316  
 Db 3864 AACCTCAATATATGCTGATACATCGTTTGGACCCCAACTATATATGCAATTTCCACAC 3923  
 Qy 4317 actctcgacagagtgatgaggtgctgtaattaaatgctggaagatgacatctgggaatg 4376  
 Db 3924 ACTCTTCCATGAGATTGATGGGCTGTGATTTAAATAGTTGACAGCATGCGCATTTGGGAAG 3983  
 Qy 4377 gagagtaagaacataatcagatgacagataatctgcttcaatctcaacttcaacaatgctt 4436  
 Db 3984 GAGAGTAAGAATAATATGATGATGACAGATTAATCTGCTTCACTTCAATATATGATT 4043  
 Qy 4437 ggcacctgtctctctcaaaagctcgacttcaactccaaaggagagtaatgctcggaga 4496  
 |||||||

Db 4044 GCCACCTGCTCTCTTCAAAAGCTGACTTCACTTCCAGGAGGAGATATGCTTGAGA 4103  
 Qy 4497 cctcaagtgataatcccaaaagatgctgcaagtggacttccaaagaacataaagatc 4556  
 Db 4104 CCTCAGGTAAATATATCCAAAGAGTGGCTGCAAGTGGACTTCCAAACATATGAAATC 4163  
 Qy 4557 acaagagtaactactcaaggaglaaaatctctgtctaccagaatgataatgaaagatc 4616  
 Db 4164 ACAGAGTAATACATCAGAGAGTAATAATCTGTCTTACAGCATGATATGAAAGATTTC 4223  
 Qy 4617 cctcatctcagaagatcaaaatgagcatagtgactctcttcttcaagaatggcaagta 4676  
 Db 4224 CTATCTCCAGAGTCAATGAATGAGCCATGATGACTCTCTTTTCAAAATGGCAAAATA 4283  
 Qy 4677 aaggttttcaaggaaataaagactccctcaactgtggtgaactctctagaccacg 4736  
 Db 4284 AAGGTTTTTCAAGGAATTAACATCTTACACACTGTGTGAATCTCTAGACCCACCG 4343  
 Qy 4737 ttaactgactgcttaacttgaatcaaccaccagaatgtggtgacagaatgtgacctgaag 4796  
 Db 4344 TTACTGACTGCTACTTGAATTCACCCCGAGAGTGGTGCACACAGATTGCCCTGAG 4403  
 Qy 4797 atggaggtctggtgctgagagcagaagacctctactactaclogagcgatc 4848  
 Db 4404 ATGAGGTTCTGTGGCTGCGAGGACAGGACCTCTACTGAGGGGTGGCACTGC 4455

RESULT 3  
 LOCUS 108644 4278 bp PAT 02-DEC-1994  
 DEFINITION Sequence 4 from Patent WO 8800831.  
 ACCESSION 108644  
 VERSION 108644.1 GI:588650  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unclassified.  
 REFERENCE 1 (bases 1 to 4278)  
 AUTHORS Pasek, M. P.  
 TITLE DNA SEQUENCES CODING FOR MODIFIED FACTOR VIII:C AND MODIFIED FACTOR VIII:C-LIKE POLYPEPTIDES AND PROCESSES FOR PRODUCING THESE POLYPEPTIDES IN HIGH YIELDS  
 JOURNAL Patent: WO 8800831-A 4 11-FEB-1988;  
 FEATURES Location/Qualifiers  
 source 1..4278  
 BASE COUNT 1244 a 942 c 946 g 1146 t  
 ORIGIN

Query Match 52.2%; Score 4148.8; DB 10; Length 4278;  
 Best Local Similarity 97.9%; Pred. No. 0;  
 Matches 4270; Conservative 0; Mismatches 2; Indels 90; Gaps 3;  
 Qy 477 gccaccagaagaatactaccctgggtgcaatggaactgtcatggaactatatagcgaatgat 536  
 Db 1 GCCACAGAAATACACTGCTGGTGCAGTGAAGTGCATGAGACTATATGCAAAAGTAT 60  
 Qy 537 ctccgtgagctgctctgtggaagcaagattctctctaagtgccaaatctttccatc 596  
 Db 61 CTCGGTGAAGCTCCGTGTGACGCAAGATTTCCTCTAAGTGCMAAAATCTTTTCCATTC 120  
 Qy 597 aacacctcagctgcttacaaaagaactggtttgtagattcaaggttccactttcaac 656  
 Db 121 AACACCTCAGTGTGTACAAAAGACTCTGTTGTAGAAATTCACGAGTACACCTTTTCAAC 180  
 Qy 657 atcgcctaaagccaaagccacctggaatgggtctgctaggttccatccatccaaagctgaagt 716  
 Db 181 ATCGCTAAACCAAGGCCACCCGTGATGGATGCTGCTAGTCTTACATCCACAGCTGAGGTT 240  
 Qy 717 tatgataagtggtcataactataagaacatggttccatctctgtcagcttccatgct 776  
 Db 241 TATGATACAGTGTGCTATTAACCTTAAGAAACATGGCTTCCATCTGTCAAGTCTTCAATGCT 300



```
OY 777 gttggtglatccctactggaaaagctctgaaggagcgtgaataatgatgcagaccagtica 836
|||||
DB 301 GTTGGTGTATCTACTGGAAGAAAGCTTCTAGGAGAGCTGATATGATGATCAGACCACTGTA 360
OY 837 agggagaaagaaagatgatataaagctctcccttggtagaagccalacalattctctgcaagtc 896
DB 361 AGGGAGAAAGAAAGATGATAAAGTCTTCCCTGGTGAAGCCATACATATGTCTGGAGGTC 420
OY 897 ctgaaagagaatggtcacaatggtcctctgcaaccactgtgcttactactacatattctc 956
DB 421 CTGAAGAGAAATGTGTCATGTGCGCTCTACCCACAGTGTCTTACTACTCATATCTTTCT 480
OY 957 catgtggaccctggttaaaagactgaatcaggccctcatctggagccctactgatatgata 1016
|||||
DB 481 CATGTGGACCTGTGTAAGAAAGACTTGATATTCAGGCTCATTTGAGCCCTCTGATATGTAGA 540
OY 1017 gaaggagctctgycacaagaaagacacagacactgtgcacaaattatatactatttgc 1076
|||||
DB 541 GAAGGAGTCTGGCCCAAGAAAGACACACACTTGACAAATTATTACTACTTTTGCT 600
OY 1077 gtaattgtatgaagggaaaagtgtgacccacagaaaaaagaactccttgaatgcagatagg 1136
|||||
DB 601 GTATTGTGTAAGGGAAGAAAGTTGGCAGTCAGAAACAAAGAACTCCTTGATGCAGGATAGG 660
OY 1137 gatgtcatctgtctcgggccctggcctaaatgycacacagtcacatggttatgtaaacag 1196
|||||
DB 661 GATGCTGATCTGCTCGGGCCCTGGCTTAAGTCACACAGTCATGATTTATGTAAACAGG 720
OY 1197 tctctg----cgaagtlctatgtatgacacagagaatcagtcatactgcatgtatgata 1253
|||||
DB 721 TCTCTGACTACAGAGTCTGATTGGATGCCAGGAATCAGTCTATGGCATGTGATGGA 780
OY 1254 atgggacacactccttgaatgacatcctaattcctcgaaggttcacacattccttgtgagg 1313
|||||
DB 781 ATGGCACACCTCTCGAAGTGCACATATCTCCGAAAGTCAACACTTTCTGTGAGG 840
OY 1314 aacctgcgaaggctctctggaatctcgcaataacttccctactgctcacaacatc 1373
|||||
DB 841 AACCTGCCAGGCGCTCTTGGAAATCTGCGCAATTACTTTCTTACGTCTCAAAACATC 900
OY 1374 ttgatgacacttggacagttcttactglttctgcatatctctccacacaacatgaltgc 1433
|||||
DB 901 TTGATGGACCTTGAGACATTTCTACTGTTTGTCTATCTCTTCCACCAACATGATGGC 960
OY 1434 atggaagcttatgtcacaagtgcagcgtctccagaaggaaaccccaactcgsaatagaataa 1493
|||||
DB 961 ATGGAAGCTTATGTCAAGTACAGAGCTGTCCAGAGAAACCCAACTGCAATGAAAT 1020
OY 1494 aatgaagaagcggaaagactatgatgatcttactgaltctgaaatgagtgtgacag 1553
|||||
DB 1021 AATGAAGAAGCGGAAGACTATGATGATGATCTTACTGATTTCTGAATATGATGTGTCAGG 1080
OY 1554 ttgatgaltgacaactcctcttcccttacttccaaatctgctcagttgcccagaagacatcct 1613
|||||
DB 1081 TTTGATGATGACAACCTCTCTTCTTATTCCAAAATTCGCTCAGTTGCCAAGAAGATCCT 1140
OY 1614 aaaacttgggtacatttaattgctgctgcaagaaggagacttggactatgctcccttaagtc 1673
|||||
DB 1141 AAAACTTGGGTACATTAATGCTGCTGAAGAAGAGACTGGACTATGCTCCCTTAATGC 1200
OY 1674 ctgcgccccgatacagaagttataaaagttacaatatgtgaacaatgqccctcagcgagtt 1733
|||||
DB 1201 CTCGCCCCCGATGACAGAAAGTTATTAAGTCAATRTTTGAACAAATGGCCCTCAGCGGATT 1260
OY 1734 ggtgaggaagtacaaaagagtcogatttaaggcatacacagatgaacacctttaagaactcgt 1793
|||||
DB 1261 GGTAGGAAGTCAAAAAGAGTCCGATTATGTGCATACACAGATGAACCTTTAAAGACTGCT 1320
OY 1794 gaagctatcagcatgaatcaggaatccttggacccttacttattggggaagttggagac 1853
|||||
DB 1321 GAAGCTATTACAGCANTGATGAGAAATCTTGGACCTTTACTTTATGGGGAAGTTGGAGAC 1380
OY 1854 aacactgttattatatttaagaatcgaagcagacaccataataacatccactcgaaga 1913
|||||
DB 1381 ACACGTGTTGATTATTTAAAGATCAACAGACACCATATTAACATTTACCTTCACGGGA 1440
OY 1914 atcaactgtatcgttcccttggatattcaaggagatbaccaaaaggtgttaaaacatttgaag 1973
|||||
DB 1441 ATCACTGATGTCCGCTTGTGATTCAGAGAGATTACCAAAAAGGTAAACATTTGAAG 1500
OY 1974 gatttccaattctgcgcagggagaatatccaataataaaygaacagtgatcgttgaagat 2033
|||||
DB 1501 GATTTTCCAAATTCGCCAGGAGAAATATTCAAAATATAATGAGACGTGACGTGAGAAAT 1560
OY 2034 gggccaactaatacagatccctcagggtgcctggaacccgctactactagttcglttaatg 2093
|||||
DB 1561 GGGCCAACTAATATCAGATCTCGGCGCTGACCCGCTTTACTGATGTTGTTATATG 1620
OY 2094 gagaagatactagcttcaagacatattggccctccctcatalctgtacaaagaatctgta 2153
|||||
DB 1621 GAGAGAGATCTAGCTTCAGAGACTAATGGCCCTCTCCATCTGCTACAAAGAAATCTGTA 1680
OY 2154 gatcaagaaggaacccagatatatgtcagacaagggaaatgtcaltccctgttctgtatt 2213
|||||
DB 1681 GATCAAGAGAGAAACCATATATGTACAGACAAGAGAAATGTCATCTGTTTCTGTATTT 1740
OY 2214 gatgagaacccgaagctgtgtacctcaagagatatacaagcttctcccaatccagct 2273
|||||
DB 1741 GATGAGAACCCAGCTGTTACTCTCAGAGAAATATACAAAGCTTTCTCCCAATCCAGCT 1800
OY 2274 ggaatgcagcttggagatccagagttccaaagctcccaacatcatatgacagatcaatgyc 2333
|||||
DB 1801 GGAGTGCAGCTTGAGAGATCCAGAGTCCAGGCTCCCAACATCATGACACAGATCAATGGC 1860
OY 2334 tatgttttggatgttggcaatgtcagtttqgttggatggagtggtgacatgttactt 2393
|||||
DB 1861 TATGTTTGTGATGATGTCAGTGTGTCACTTGTGTTGATGAGTGGCATPACTGTGTACTTT 1920
OY 2394 ctaagcaattggagacagaacagactccttctcttctctctctctctctctctctctca 2453
|||||
DB 1921 CTAAGCATTTGAGACAGACAGACTTCTTCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCAA 1980
OY 2454 cacaanaatgltctatgaagacacactcaacctatcccatctcagaagaaactgtcttc 2513
|||||
DB 1981 CACAAAATGCTTATGAAGACACACTCCACTATTCCTCATTCAGGAGAAAGTCTTTC 2040
OY 2514 atgtcgtatggaaaacccaaggtctatgtatcttgggtgtgcccacaactcgaacttccgaac 2573
|||||
DB 2041 ATGTGATGGAAGAAACCAAGTCTATGATTTGTGGGTGCCAACCTCAGACTTTCGGAAC 2100
OY 2574 agagcatgaccgccttactgaagttctagttgtgacaagaacactggtgatltatc 2633
|||||
DB 2101 AGAGCATGACCGCTTACTGAAGTTTCTAGTTGTGAACAAGAACACTGGTGAATTTATAC 2160
OY 2634 gaggacagttatgaagatattcagacacttgcgtgagttaaaaaacaaatgcattgaaaca 2693
|||||
DB 2161 GAGACAGTTTATGAAGATATTTCAGCATACTTCTGTGATTAATAAATGCAATGATGAACCA 2220
OY 2694 agaagcttctcccaagaattcaagacacctagccttgagcaaaagcaatttaatgccaac 2753
|||||
DB 2221 AGA----- 2223
OY 2754 ccaccagtlcttgaacgcgalcaacggaataaactcgtlactacttctcagtcagatcaa 2813
|||||
DB 2224 -----GAATTAACCTGCTACTTACTCTTCACTGATGATCA 2256
OY 2814 gaggaatltgactatgatatlacatatcagttgaatltgaagaagaaatlttgacatt 2873
|||||
DB 2257 GAGCAAAATTTGACTATGATGATACATATCAAGTGAATGAAGAAGAAATTTGACATT 2316
OY 2874 tatgatgaagltgaanaatcagcccccgagcttcttaaaagaaacagacatatatt 2933
|||||
DB 2317 TATGATGAGGTGAANAATCAAGCCCCGAGCTTTTCAAAAGAAACCGACACTATTTT 2376
OY 2934 atgtcgtcagtgagagagctctggtattatgtatgtatgctcccaactgttcttaaga 2993
```



|||||  
Db 2377 ATTGCTGAGTGGAGAGGCTCTGGGATTATGGATGAGTAGCTCCCAATGTTCTAATA 2436  
OY aacagggctcagagtgagagtgccctcagtcacaagaattgttttcaggaattact 3053  
Db 2437 AACAGGGCTCAGAGTGGAGTGTCCCTCAGTTCAGAGAAAGTTGTTTCCAGGAATTACT 2496  
OY gatggctccttactcagccctataccgttgagaaactaaatgaacatttggagctcctg 3113  
Db 2497 GATGGCTCCTTACTCAGCCCTTATACCGTGGAGAACTAATGAACATTGGAGCTCTCG 2556  
OY 3114 gggccatatataagagcagaagtltgaagataatcatatcatgttaacttcagaaatcagcc 3173  
Db 2557 GGGCCATATATAAGAGCAGAAAGTTGAAGTAATCATGTGTAACCTTTCAGAAATCAGGCC 2616  
OY 3174 tctcgtccctatctccttactatctagcccttacttcttaagaaagatcagaagcaga 3233  
Db 2617 TCTGCTCCCTATCTCTTCTCATCTCAGCCCTTATTTCTTATGAGAGAAATCAGAGCGAAGA 2676  
OY 3234 gcagaacctagaaaaaactltgtcaagcctaaatgaaccaaaacttacttctggaagtgt 3293  
Db 2677 GCAGAACTAGAAAAAAGCTTGTCAAGCCTAATGAACCAAAAGCTTCTTGGAAAGTG 2736  
OY 3294 caaatcatataggaccacactaaagaatgagttgacttgaagaagcctgagcttattctct 3353  
Db 2737 CAACATCATATAGCACACCACATAAGATGAGTTGACTGCAAGCCCTGGGCTTATTTCTCT 2796  
OY 3354 gatctgaacctgaaaaaagtgtgcactcagagccttgatcgaccctctctgtctcagc 3413  
Db 2797 GATGTTGACCTGGAAAAAGATGTGCACCTCAGGCTGATTTGAGCCCTTGTGCTGCTCCAC 2856  
OY 3414 actaacacacagaaacctgtgtcaatggagaaagtgaacagtaacgaattgtctgtt 3473  
Db 2857 ACTAACACACACTGAACCCCTGCTCATGTGGAGACAAAGTGACAGTAACGAATTTCTGTTT 2916  
OY 3474 C---tcacatcttctgtagagaccaaagctgtgtacttcaatgaaatltgaaagaac 3530  
Db 2917 TTCCCTACACTCTTGTGATGAGACCAAAAGCTGTACTTCAGTGAATATGGAAGAAAC 2976  
OY 3531 tgcagggctccctcgaataatcagagatgagaatcccaactttaaagaatatacgtctc 3590  
Db 2977 TGCAGGGCTCCCTCGAATATCCAGATGGAAGATCCCACTTTTAAAGAGATTATCCCTTC 3036  
OY 3591 catgcaatcaatggtcacaataatggaatacactggtctgtgaatggcacaagataca 3650  
Db 3037 CATGCAATCAATGGCTACATATGATGATCACTACCTGCTGATGAGCTCAGAGATCAA 3096  
OY 3651 aggaattgattgattctcctcagagatggagcaatgaatgaatcatctcttcttctc 3710  
Db 3097 AGGATTCAGATGATATCTGCTCAGCATGGGACCAATGAAGAACATCCATTCTATTCTATTC 3156  
OY 3711 agtgaacatgtgtcactcgtacgaaaaaaagagagatataaaatggtcacgttacaatcc 3770  
Db 3157 AGTGAGACATGTGTCTACTGTACGAAAAAAGAGAGATATAAATGGCACTGTATCAATCTC 3216  
OY 3771 tatccaggtgttttctgagaacagtggaatgttacaatccaaagctggaaatttggcggtg 3830  
Db 3217 TATCCAGGTGTGTTTGAAGACAGTGAAGATGTACCATCCAAAGCTGGAAATTTGGCGGTG 3276  
OY 3831 gaatgacctatttggagacatcatatcgttggatgagacaccttcttctgtgtacacg 3890  
Db 3277 GAATGCCCTTATTTGGAGACATCTACATGCTGGATGAGACACTTTTCTGTGCTGACACG 3336  
OY 3891 aataagtgtcagacctcccttggaaatgtctcttgacacatlagagatttccagattaca 3950  
Db 3337 AATAAGTGTCAAGACCTCCCTGGAGATGGTCTTGACACATATAGAGATTTCAGATTTAA 3396  
OY 3951 gctcagaagaacatcagacagtgggcccaaaagcttggccaagactcatatctcagatca 4010  
Db 3397 GCTTCAGAGACATATAGACAGTGGGCCCCAAAGCTGGCCAGACTTCATTATTCGGATCA 3456  
OY 4011 atcaatgcttgagagaccaaagagcccttcttctgtatcaagtgatctgttggcacca 4070  
|||||

Db 3457 ATCAATGCTGGAGACCAAGAGACCCCTTCTTGGAATCAAGTGATCTGTTGGACCA 3516  
OY 4071 atgattatcacggacatacagaaccaggtgtcccgltcagaagtctccagccttatactc 4130  
Db 3517 ATGATTATTACGGGATTCAGACAGCCAGGCTGCCCTCAGAAATTTCTCCAGCTCTACATC 3576  
OY 4131 tctcagltatcatcatatgatatagcttgaatggaaagtggtcagacttatacgaagaat 4190  
Db 3577 TCTCACTTATCATCATGTATGATGCTTTCATGGGAAGAGTGGCAGACATTAATCGAGAAAT 3636  
OY 4191 tccagtgaaacttaatgtcttcttcttggcaatgtggatltatctggatataaacacat 4250  
Db 3637 TCCACTGGAACTTAAATGCTTCTTGGCAATGTGGATTCAATGGGATTAACACAT 3696  
OY 4251 attttaaccctccaaltcttgcatalacacgttgcacccaacatlatagaalt 4310  
Db 3697 ATTTTAACTCCATATTTGCTCGATTCATCCGTTTGACCCAAACATTAATATGCAATT 3756  
OY 4311 cgcagacatcttcgcatgtgagttgagtggtgtgatttaaatagltgcagatgcattg 4370  
Db 3757 CCGAGCACTCTTCGATGAGATGGCTGTGATTTAAATAGTTGCAGCATGCCATTG 3816  
OY 4371 ggaatggagatfaaagcaatcatcagatgacagatgactgtcttacttacttacaat 4430  
Db 3817 GGAATGAGATGAAGCATATACATACATGACAGATTACTGCTTACTCTTATACAT 3876  
OY 4431 atgttgcacactgtgtctctctcctcaaaagctgcacttcaactccaagggagagtaatgcc 4490  
Db 3877 ATGTTGGCACCTGGTCTCTTCAAAAGCTGACTTCACTCACTCACTTATACAT 3936  
OY 4491 tggagaacctcaggtgaaataatccaaaagatgtgtcgaatgtgaccttccagaagaatg 4550  
Db 3937 TGGAGACCTCAGAGTGAATATCAAAAGAGTGGCTGCAAGTGGACTTCCAAAGCAATG 3996  
OY 4551 aaagtacagaggttaactcactcagagagatgaatctctgtcttacaagatlatgtgaag 4610  
Db 3997 AAAGTACAGAGATACATCTAGGAGATGAATCTGCTTACAGCATATATGTGAAG 4056  
OY 4611 gatttccatctcactcagcagatcaagaatgagcactcagtgactctcttcttccaagaatggc 4670  
Db 4057 GAGTTCCTCATCTCCAGCAGATCAAGATGGCCATCAGTGGAGCTCTTTTTCAGAAATGCC 4116  
OY 4671 aaagtaaaagtlttccaaggaaatcaagaacaccttcaacactgtgtgtaactcttagac 4730  
Db 4117 AAAGTAAAGTGTTCAGGAAATCAAGACCTCTTACACACCTGTGTAACCTCTTACAC 4176  
OY 4731 ccaacgttactcagctcagcttcaatctcaaccccaagatgtgtgtcaccagatggc 4790  
Db 4177 CCACCGTTACTACCTGCTACTTGAATTCACCCCAAGATGGGTGCACACATTTGCC 4236  
OY 4791 ctgagatgagagttctgtgctgcagaggacagacctctac 4832  
Db 4237 CTGAGGATGAGAGTTCTGTGGCTGCGAGGCACAGACCTCTAC 4278  
|||||

RESULT 4  
LOCUS 108643 4281 bp  
DEFINITION Sequence 3 from Patent WO 8800831. PAT 02-DEC-1994  
ACCESSION 108643  
VERSION 108643.1 GI:588649  
KEYWORDS  
SOURCE  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 4281)  
AUTHORS Pasek,M.P.  
TITLE DNA SEQUENCES CODING FOR MODIFIED FACTOR VIII:C AND MODIFIED FACTOR VIII:C-LIKE POLYPEPTIDES AND PROCESSES FOR PRODUCING THESE  
JOURNAL POLYPEPTIDES IN HIGH YIELDS  
FEATURES Patent: WO 8800831-A 3 11-FEB-1988;  
Source Location/Qualifiers  
1..4281



BASE COUNT 1245 a 943 c 946 g 1147 t  
ORIGIN

Query Match 52.28; Score 4147.2; DB 10; Length 4281;  
Best Local Similarity 97.9%; Pred. No. 0;  
Matches 4269; Conservative 0; Mismatches 3; Indels 90; Gaps 3;

QY 477 gccacccgaagaatctactctggtgctgagtggaactgtcatgtggaactatataatgcaagatgat 536  
|||||  
DB 4 gccacccgaagaatctactctggtgctgagtggaactgtcatgtggaactatataatgcaagatgat 63  
QY 537 ctgagtgagcgcctgtggaagcaagatttctctctagagtgccaaaatctcttccattc 596  
|||||  
DB 64 ctgagtgagcgcctgtggaagcaagatttctctctagagtgccaaaatctcttccattc 123  
QY 597 aacacccatagtcgtgtcaaaaaaagactctgttggtagaattcaagtggtcaccttccaac 656  
|||||  
DB 124 AACACCTCAGTCGTGTAACAAAAGACTCTGTTGTAGAAATTCACGAGATCACCTTTTCAAC 183  
QY 657 atgcctaaagcaagggccacccctggaatggtctgctgaagtcctaccatccaggtgaagtt 716  
|||||  
DB 184 ATCGCTAAGCCAAAGCCACCTGGATGGGTCTCTAGGTCTTACCATTCCAGGCTGAGGTT 243  
QY 717 tatgatacaagtgtgctattacacttaagaacaatggtctccatctctgcaagcttcatgct 776  
|||||  
DB 244 TATGATACAGTGGTCAATTAACCTTAAGAACATGAGCTTCCATCTGTGACGCTTTCATGCT 303  
QY 777 gttgtgtatctactctggaagactctctgaaggagctgaataatgaatgaatgaatgaatgaat 836  
|||||  
DB 304 GTTGCTGATCTCTAGTGAAGGCTTCTGAGGAGCTGATATGATGATGATGATGATGATGATGAT 363  
QY 837 agggagaagaagaatgataaagcttccctgggtggaagccaatataatgcttgcagagtc 896  
|||||  
DB 364 AGGAGGAAGAAAGATGATAAAGTCTCCCTGGTGAAGCCATACATATGCTGACAGGTC 423  
QY 897 ctgaaagaagaatgctcaatgctctgcaaccactgtgcttactactactactactactcttct 956  
|||||  
DB 424 CTGAAAGAGATGCTCCAAATGCTCTGACCCACTGTGCTTACTACTACTACTACTACTACTACT 483  
QY 957 catgtgacctgtgtaaaagactgaattcaggcctcatctggagccctactagtaatgata 1016  
|||||  
DB 484 CATGTGACCTGGTAAAGACTTGAATTCAGGCTCATTTGGAGCCCTACTATGATGATGATA 543  
QY 1017 gaaggaggtctggtccaaaggaagaacacagacctgtgacaaatattactactcttctgtct 1076  
|||||  
DB 544 GAAAGGAGTGTGGCCAGGAAAGACACACACCTTGCACAAATTTATATACCTTTTGTCT 603  
QY 1077 gtaattgaagggaagaatgtgacctcctcagaagaagaagactccttgatgacagatagg 1136  
|||||  
DB 604 GTATTATATGAAGGAAAGATGAGACTCAGAAACAAAGAACTCTTGATGACAGGATAGG 663  
QY 1137 gatgtgcatctgtctgggctgtgacctaaatgacacagtcacatgg tlaatgaacagg 1196  
|||||  
DB 664 GATGCTCATCTGCTGGGCTGGGCTTAAATGCAACACAGTCAATGGTTATGTAACAGG 723  
QY 1197 tctctg---ccaggctcgattggaatgcacaaagaatcagtcataatggaatggaatgga 1253  
|||||  
DB 724 TCTCTGTACCAAGTCTGATTTGATGACACAGGAAATCAGTCTATTTGGCATGTGATTTGA 783  
QY 1254 atggagacacactcctgaaggacactaataatctctggaagtgacacattctctgttgaag 1313  
|||||  
DB 784 ATGGGACACACTCTCTGAAGTCACTCAATATCTCTCAAGGTCAACAAATTTCTGTGAGG 843  
QY 1314 aacacatgcgaagcgctcttggaatctgcacaa taaacttctctactgtcacaacatc 1373  
|||||  
DB 844 AACCATGCGCAGGCGCTCTGGAATCTCGGCATTAACCTTCTTACTGCTGCAAAACATCTC 903  
QY 1374 ttgatgaccttggaaggttctactgttttctgtcatatctcttcccaacaatgatgagc 1433  
|||||  
DB 904 TTGATGAGACCTTGACAGTTTCTACTGTTTGTTCATATCTCTTCCCAACAAATGATGGC 963

QY 1434 atggaagcttatgtcaaaagtacagacgctgtccagagaagaccccaactacgaatgaaaaat 1493  
|||||  
DB 964 ATGGAGCTTATGTCAAAAGTAGACAGCTGTCCAGAGAACCCCACTCGAATGAAAAAT 1023  
QY 1494 aatgaagaagcgggaagactatgatgatgatactactgattctgaaatggaatggtgacag 1553  
|||||  
DB 1024 AATGAGAGAGCGGAAGACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1083  
QY 1554 ttgatgatgacactctctctctctctctctctctctctctctctctctctctctctctctctctct 1613  
|||||  
DB 1084 TTTGATGATGACAACTCTCTCTCTTATTCAAATTCGCTCAGTTGCTCAAAACATCTCT 1143  
QY 1614 aaaacttgggtacattacatgctgcgtgaagaagagacgtggagactatgctcccttactc 1673  
|||||  
DB 1144 AAAACTTGGTACATTTACATTTGCTGTAAGAGAGAGACTGGACTATGCTTCCCTTACTC 1203  
QY 1674 ctgcgccccgaatgacagaagttataaaagtacataattgaaacaatggtccctcaagatgt 1733  
|||||  
DB 1204 CTGCCCCCGATGACAGAGATTATAAAGTCAATATTGAAACAATGGCCCTCAGGGGATT 1263  
QY 1734 ggtaggaagtacaaaaaagtcgattatggtacatacacagatgaaccttgaagctgt 1793  
|||||  
DB 1264 GGTAGGAAGTCAAAAAAGTCCGATTTATGGCATACACAGATGAACCTTTAAGACTCGT 1323  
QY 1794 gaagctatcagacatgaatcagaagacttggagacttacttcttggggaagttggaagac 1853  
|||||  
DB 1324 GAAGCTATTCAGCATGATGATGAGAACTTGGAGCTTACTTATGAGGAAGTTGGAGAC 1383  
QY 1854 acaactgtgatatataatlaaagaatcaagcaagcagacacataatacatctaccccaagca 1913  
|||||  
DB 1384 ACACGTGTTGATTATATTTAAACAATCAAGCAGACAGACATFMAACATCTACCTCCACAGCA 1443  
QY 1914 atcaactgattgcgctcctctgttattcaagaagatccaaaagtgtaaaacattggaag 1973  
|||||  
DB 1444 ATCACTGATGTCCTCTCTTGTATTCAGGAGATTACCAAAAGGTGTAATAACATTTGAG 1503  
QY 1974 gatttccaatctgtccagaggaataatcaatataaagtgaagcagtgatgaagat 2033  
|||||  
DB 1504 GATTTTCCAATTCGCCAGAGGAATATTCAAAATATAATGAAAGTACGTGTAAGAT 1563  
QY 2034 gggcacaactaaatcaagatcctcgtgtccctgacccgctactactactctctctctctctct 2093  
|||||  
DB 1564 GGGCCAATAATATAGATATCCCGGTGCTGACCCGCTATTCCTAGTTGTTGTTAAATG 1623  
QY 2094 gaagagatctaggttcaagagctcattggccctctctctcaatctgctataaagaatctgta 2153  
|||||  
DB 1624 GAGAGAGATCTAGCTTCAAGACTCATTTGGCCCTTCCATCTCTCAAAAGAAATCTGTA 1683  
QY 2154 gatcaagaagaagaacagataatgtcagaagaaggaatgcatccgttctctgtatct 2213  
|||||  
DB 1684 GATCAAGAGAGAAACAGATATATGTACAGCAAGAGAAATGTATCATCTGTTTCTGTATTT 1743  
QY 2214 gatgagaacccaagctgtgtacctcacaagaagaatataacaagcttctctcccaatccagct 2273  
|||||  
DB 1744 GATGAGAACCGAAGCTGTATCTCACAAGAGAAATATACAGCTTTCGCCCAATCCAGCT 1803  
QY 2274 ggaagtcagcttggaagatccaagagttccaagctccaacatcatgacaagaatcagatgagc 2333  
|||||  
DB 1804 GGAAGTGAAGCTTGAAGATCCAGAGATTCCAAGCTTCAACATCATGACAGCATCAATAGGCT 1863  
QY 2334 tatgttttgaatagttgagagtgctaggttggattggttgaatgagtgagatctgttaact 2393  
|||||  
DB 1864 TATGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1933  
QY 2394 ctaagcatlgtgaagacagactgaacttctctctctctctctctctctctctctctctctctctct 2453  
|||||  
DB 1924 CTAGCATTTGGAGACAGACAGACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1983  
QY 2454 cacaaaatggtctatgaagacacactacccatctctcaatctcaggaagaacatgctctc 2513  
|||||  
DB 1984 CACAAATGTGCTATGAAAGACACACTCACCTTATTCCTATTCCTCAGAGAAAGTGTCTTC 2043  
QY 2514 atgctcgttgtaaaacccaggttctatgattctctgggtggtcccaactcagacttccggaac 2573



```

Db 2044 ATGTCGATGGAAAAACCCAGGCTCTATGGATTCTGGGGTGGCCACAACTCAACACTCTTTCGGAAC 2103
QY 2574 agagcatgacgcgccttactgaaggtttctagttgtgacaagaacacgtgagatattac 2633
Db 2104 AGAGGCAATGACCGCTTACTGAAGTTTCTAGTTGTGACAAACACACTGCTGTTATTTAC 2163
QY 2634 gaggacagttatgaagatattcaagacatactgtctgagtaaaaaaactgccaatgaacca 2693
Db 2164 GAGGACAGTTATGAAGATATTTCACCATACTTGTGAGTAAAAAACATGCCATTGAAACA 2223
QY 2694 agaaagcttccccaagaattcaagacacccctagcactagcgaaagcaatttaatgcacc 2753
Db 2224 AGA----- 2226
QY 2754 ccaccagcttgaagcgcatacgaacggaataactcgtactactccttcagtcagatcaa 2813
Db 2227 -----GAAATACTCGTACTACTCTTTCAGTCAGATCAA 2259
QY 2814 gaggaaatlgactatgatalacatacagttgaatlgaaagaaagaaatltgacat 2873
Db 2260 GAGGAAATTTGACTATGATGATACCATATCAGTTGAAATGAAGAAGAGATTTTGACATT 2319
QY 2874 tatgataagagttgaanaatlcagagaccccgacgcttccaaaagaaaaacagacacatttt 2933
Db 2320 TATGATGAGATGAAAAATCAGAGCCCGCACCTTTCAAAAAACAACAGACATATTTT 2379
QY 2934 atgtcgtgagtgagagagctcgtgagattatgagatgagtagctccacagatlttcaaga 2993
Db 2380 ATTGCTGAGTGGAGAGGCTCTGGGATTTATGGATGAGTAGCTCCGCCACATGTTCTAAGA 2439
QY 2994 aacagggctcagagtgagtgagtgccctcagttcaagaagaatgtgttccagaatltact 3053
Db 2440 AACAGGCTCAGAGTGGAGTGCCTCCCTCAGTTCAGAAAGTGTTCACGAAATTTACT 2499
QY 3054 gatgctcccttactcagcccttataccgttgagagactaaatgaacatttggactccg 3113
Db 2500 GATGGCTCTTACTACACCCCTTATACGTGAGAACTPAATGAACATTTGGAGCTCCG 2559
QY 3114 gggcatalataaagagcgaagttgaagaataalacatgtagtaacttcagaatcagcc 3173
Db 2560 GGGGCATATATAGAGCAGGAAGTTGAAGATATATCATGTAACTTTCGAATCAGGCC 2619
QY 3174 tctgtccctatcccttactatctagccttattctatlgagaaagatcagaagga 3233
Db 2620 TCTGTGCTCTTCTCTCTTATTTCTAGCCTTATTTCTATGAGAGATCAGAGCAAGA 2679
QY 3234 gcgaagaactgaaaaaacttgcgaagcctaatagaaccaaacttactttggaaagtg 3293
Db 2680 GCAGAACCTTAAAAAACTTGTCAAGCTTAAATGAACCAAAAACCTTCTTGGAAAGTG 2739
QY 3294 caacatcatatgacaccacaaagaatagtttgaactgcaaaagccttgagctatttctc 3353
Db 2740 CAATATCATATAGGACCCACATAAAGATAGATTGTGACTGCAAAAGCCTGGCTTATTTCT 2799
QY 3354 gatgtgaactggaanaaagatgtgacactcagcctgatttgaccccttcgtgtgcac 3413
Db 2800 GATTTTACCTGGAAAAAGATGTGCACCTCAGGCCCTGATTTGGACCCCTCTGCTGCCAC 2859
QY 3414 actaagaacatgaacccctgcatactgagagacaagtgacagatagaagaatttgcctgttt 3473
Db 2860 ACTAACACACTGAACCCGCTCATGGGAGACAAAGTACAGTAAGAAATTTGCTCTGTTT 2919
QY 3474 t-----tcaacatcttltgatagaaccaaaagctlgtagtctcaactgaaataatlgaaagaaac 3530
Db 2920 TTCTCTACCATCTTTGATGAGAGACCAAAAGCTGCTACTTACAGTGAATAATGGAAGAAAC 2979
QY 3531 tgaagagctccctgcataatccagatggaagatcccaactttaaagagaattatcgcttc 3590
Db 2980 TGCAGGGCTCCCTGCAATATCCAGATGGAAGATCCCATTTTAAAGAGCAATATATGCTTC 3039
QY 3591 catgcaatcaatgtgctataatagtatactactgcttagtaatgtcttaagatcaa 3650

```

---

```

Db 3040 CATGCAATCATGCTACTAATAATGATACACTACCTGGCTTAGTAAATGCTCAGACATCA 3099
QY 3651 aggattcgaatggatctgtccagcatlgggcaagcaatgaanaacatccattcatcatttc 3710
Db 3100 AGGATTGCAATGGTATCTGCTCAGCATGGGCACACATGAAACATCATTTATTCATTTC 3159
QY 3711 agtggacatgtgtcaactgtacgaaaaaagaagagatlaaaatggtgcactgacaatc 3770
Db 3160 AGTGACATGTTGCTCAGTACGTAAGAAAAAGAGAGTATTAATAATGSCACTGTACAATCTC 3219
QY 3771 tatccagtgcttttgagacagtggaatgyltaccataccaagcgtgaatttgcgggtg 3830
Db 3220 TATCCAGGTGTTTTGAGACATGTGAAATGTTTACCATCTCAAGCTGGAATTTGGCGGTTG 3279
QY 3831 gaatgccttatgtggagacatcatalatgctggatgagacacatttctgtgtgacagc 3890
Db 3280 GAATGCTTATTTGGGAGAGATCTACATGCTGGATGAGACACATTTTCTGTGTACAGC 3339
QY 3891 aataagtgtaagactcccttggaaatggtctctggaacatlagagatlttcaagtaca 3950
Db 3340 AATAAGTGTCAAGACTCCCTGGGAATGGCTCTGACACATTAAGATTTTCAGATTACA 3399
QY 3951 gcttgaagaacaatlgagacagtgggcccaaaagctggccgaactcatctatccggatca 4010
Db 3400 GCTTCAGACAAATATGAGAGTGGCCCAAAAGCTGGCCAGACTTATTTCCGAGTCA 3459
QY 4011 atcaatlgccttgagacaccaagagagcccttctcttgatcaaaagtgtgactgttgacaca 4070
Db 3460 ATCAATGCTGTGAGCACCAAGAGGCCCTTTCTTGATCAAGTGTGATCTTGTGGACCA 3519
QY 4071 atgattatlcacgcatcaagaaccagagtgcccgtaagaagtttccagccttcaatc 4130
Db 3520 ATGATTATTCAGGCAATCAAGACCAGGATGCCCTCCAGAAAGTCTCCAGCCTCTACATC 3579
QY 4131 tctcagttatcaatgatatgattgtcttgatggaaagagtgagagactatcagaggaat 4190
Db 3580 TCTCAGTTTATATGATGATGATGATCTTGTGATGGAGAAAGTGGACACTTATTCAGAGAAAT 3639
QY 4191 tccacttgaaccttaatgtgtcttcttcttggaatgtggaatcaatctlggaataaaacaaat 4250
Db 3640 TCCACTGGAACCTTATGAGTCTTTTGGCAATGTGATCATCTGGATTAACACAT 3699
QY 4251 attttaaccctccaaattatgtctgatacatccglttgcaaccaatcatlataagat 4310
Db 3700 ATTTTAAACCTTCCAATATTGCTGATACATCCCTTGGCAACCAACTTAATTAAGATT 3759
QY 4311 gcgaagcacttcggaatgagttgagtggtgtgtgattaaatgttgaagaatccatg 4370
Db 3760 CCGAGCACTCTTCGATGGAGTGTGATGGGCTGTGATTAATAGTTTGCAGCATGCCATTG 3819
QY 4371 ggaatlgagagtaaaagcaaalatacgaatgcacagatlaactgtlcaatccatcttacaat 4430
Db 3820 GGAATGGAAGTAAGCAATATACATGACAGATTACTGCTCATCTTACTTTTAAACAAT 3879
QY 4431 atgtttgcaactgtgtctccttcaaaagctggaactcaactcccaaggaggaatgaatgc 4490
Db 3880 ATGTTTGCACCTGTGCTCTTCAAAAGCTGACACTTCACCTCAAGGAGAGATTAATGCC 3939
QY 4491 tggagacctcgaatgaataatccaaaagatgtgctggaagtgaggacttccagaagaatg 4550
Db 3940 TGGAGACCTCAAGTGAATATATCAAAAGAGTGCTGCAAGTGGACTTCAGAAAGCAATG 3999
QY 4551 aaagtacagagtaactactcaaggagtaaaatctcgtcttaccagcatgtatgtgaag 4610
Db 4000 AAAGTCACAGAGTAATACATCAGCGAGTAATAATCTGCTTACAGAGATATATGTGAAG 4059
QY 4611 gagttccatctccagcagatcaagaatggccaatcagtggaactccttcttcaagaatggc 4670
Db 4060 GAGTTCCTCATCTCCAGAGTCATMAATGATGATGATGATGATGATGATGATGATGATGATG 4119
QY 4671 aaagtaaaagtttccaggaataaagaactccatcaaacctgtgtggaactctcagac 4730
Db 4120 AAAGTAAGGTTTTTCAAGGAATCAAGACTCTTCAACACTGTGTGAACTCTCTAGAC 4179

```



QY 4731 ccacccgttactgcgtcgtctaccccttcgaatccaccccaaggttcgggtgcaccagattgccc 4790  
|||||  
Db 4180 CCACCGTACTGACTCGCTACTCTTCGATTCACCCCAAGTTGGGTGCGCACGATTTGCC 4239  
QY 4791 ctgaagatggaggttcctgggtcgcagagcaagacctctac 4832  
|||||  
Db 4240 CTGAGATGGAGGTCTGGGCTCGAGGCGACAGACCTCTTAC 4281  
RESULT 5  
LOCUS 108642 4548 bp PAT 02-DEC-1994  
DEFINITION Sequence 2 from Patent WO 8800831.  
ACCESSION 108642  
VERSION 108642.1 GI:588648  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 4548)  
AUTHORS Unclassified.  
TITLE DNA SEQUENCES CODING FOR MODIFIED FACTOR VIII:C AND MODIFIED FACTOR  
VIII:C-LIKE POLYPEPTIDES AND PROCESSES FOR PRODUCING THESE  
POLYPEPTIDES IN HIGH YIELDS  
JOURNAL Patent: WO 8800831-A 2 11-FEB-1988;  
FEATURES Location/Qualifiers  
source 1..4548  
BASE COUNT 1344 a 1004 c 1002 g 1197 t 1 others  
ORIGIN

Query Match 51.4%; Score 4085; DB 10; Length 4548;  
Best Local Similarity 95.1%; Pred. No. 0;  
Matches 4325; Conservative 0; Mismatches 31; Indels 192; Gaps 3;

QY 477 gccaccagaagatctactacctgaggtgcagtggaactgtcatggagcatatgcaagtgat 536  
1 GCCACCAAGAAAGTACTACTCGGTGCGAGTGGAACTGTGATGGACTATATGCAAAAGTGAT 60  
Db 1 GCCACCAAGAAAGTACTACTCGGTGCGAGTGGAACTGTGATGGACTATATGCAAAAGTGAT 60  
QY 537 ctccgtgagctgcctgttgagcgaagaattcctcctagaagtgccaaatcttccatlc 596  
|||||  
Db 61 CTCGGAGAGTGGCTGTGGAGCGCAAGATTTCCCTAGAGTGCCAAATCTTTCCATTTC 120  
QY 597 aacacctcagtggtgtacaagaagactcgtttgtagaattcaagggttcaacctttcaac 656  
|||||  
Db 121 AACACCTCAGTCTGTACAAAAGAGCTGTGTGTAGAAATTCACGGATCACCTTTTCAAC 180  
QY 657 atcgccaagccaagggccacctgagtggtgtcgttagtctcctaccatccagggcttgaggt 716  
|||||  
Db 181 ATGGCTAAGCCAAAGGCCACCTGATGGGTGCTCTAGGTCCTACATCCAGGCTGAGGTT 240  
QY 717 tatgatcagttggtlcatltaoacttaagaacatggtctccatcctgtlcaagttctcatgct 776  
241 TATGATACAGTGTCTACACTTAAGAACATGCTCCATCCATCTCGAGTCTTCATGCT 300  
QY 777 gttgtgtatctctactcgtgaagactcttgaggaagctgtaatagtatcagacagctcaa 836  
|||||  
Db 301 GTTGGGTATCTCTACTACGAAAGCTTCTGAGGGAGCTGAATATGATATCAGACCACTCA 360  
QY 837 agggagaagaagatgataaagcttcctccctggtgtggaagccatacatatgtctgagagtc 896  
|||||  
Db 361 AGGGAGAAGAAGATATATAAGTCTTCCTCGTGGAGAACCATATATGCTTGCGAGGTC 420  
QY 897 ctgaagaagaatggtccaatggtccttgacccactgtgccttactactactaaltcttct 956  
|||||  
Db 421 CTGAAGAAGATGCTCAATGAGGCTCTGAGCCACACTGTGCTTACTACTACATATCTTTCT 480  
QY 957 catgtgacctgtgtaaagaacttgtaattcaagcctcaattggagcctactagatagtaga 1016  
|||||  
Db 481 CATGTGAGACTGTGTAATAAGACTTTGAATTCAGGCTCATTTGGACCTCTACTAGTATATAG 540

QY 1017 gaaggagtcctgcgccaaggaagaacacagaccttgcacaatttatactacttttgc 1076  
|||||  
Db 541 GAAGGAGTGTGGCCAAAGGAAGAACACACACTTGACAAATTTATACACTTTTTCGT 600  
QY 1077 gtaattgataaaggaaagttggtcaactcagaacaagaactccttgaatgcagatag 1136  
601 GTATTGATGAAGGAAGAAAGTTGGCACTCAGAAACAAAGAACTCTTGATGACGAGTAGG 660  
QY 1137 gatgtgcactgtcgtcgggcttggtcctaaatgcacacagtcgaatggttatgtaaacagg 1196  
661 GATGCTGCATCTCTCGGGCTTGCCCTTAAATGCACACACTCATAGTTATGTAAACAGG 720  
QY 1197 tctctg---ccaggtcgtgattgattccacaaggaataatcagttcattggcattgattgga 1253  
|||||  
Db 721 TCTCTGCTACCAAGTCTGATTGGATGCCACAGAAATCACTTATTTGGCATGTGATTGGA 780  
QY 1254 atggcaccacatcctctgaagtgacatcaatalctcctcgaagtgacacatcttcttgagg 1313  
781 ATGGCACCACTCCTGAAAGTGCACATCATATTCCTCGAAGGTGCACACATTTCTTGAGG 840  
QY 1314 aaccatcgccagcgctccttggaatctgcgcaataacttccctactctcaaacatc 1373  
|||||  
Db 841 AACCATCGCCAGCGCTCTTGGAATCTCGCAATACTTCTCTACTGCTCAAAACCTC 900  
QY 1374 ttgatgaccttggaagtttctactgtttgtgcatactcttcccaacaatgattggtc 1433  
901 TTGATGACCTTGACAGTTCCTTACTGTTTGTGTCATATCTCTTCCACCAACATGATGGC 960  
QY 1434 atggaagcttaigtcaaaagtagacagctgtccagaaggaaccccaactagatgaaat 1493  
961 ATGGAAGCTTATGTCAAAAGTAGACAGCTGCCAGAGAACCCCACTACGAATGAATAAT 1020  
QY 1494 aatgaagaagcgaagaactctatgattgattctactgattctgaatagtgtggtcagg 1553  
|||||  
Db 1021 AATGAAGAAACGGGAAGCATATGATGATGATCTTACTGATTTCTAAAGATGGTGTCAGG 1080  
QY 1554 ttgatgatcaaacctctcctccttatccaaattgcgtcagttgcgaagaagatcct 1613  
1081 TTGATGATGATCAACCTCTCTCTCTTATTCCAATTCGCTCATTTGGCAAGAACATCCT 1140  
QY 1614 aaaacttgggtacattacattgtcgtcgaaggaagagacttggactatgtcccttagtc 1673  
1141 AAAACTTGGGTACATTATACATTTGCTGGAAGAGAGAGACTGGGACATATGCTCTTATGTC 1200  
Db 1141 AAAACTTGGGTACATTATACATTTGCTGGAAGAGAGAGACTGGGACATATGCTCTTATGTC 1200  
QY 1674 ctgcgcccgatgacagaagtataaaatgcaatatttgacaatggtccctcagcgatt 1733  
1201 CTGCCCCCGATGACGAAGATTTAAAGCAATATTTGAACAATATGCCCCCTCAGCGGATT 1260  
QY 1734 ggtgaggaagtacaaaaaagtcgcatattatggcaatacaagaatgaacacttgaagactcgt 1793  
1261 GGTAGGAAGTACAAAAAAGTCCGATTTATGCGATACACAGATGAACCTTTAAGACTCGT 1320  
QY 1794 gaagcattcgaatgaatcgaagacttgggaacttacttataatgggaagtggagac 1853  
1321 GAAGCTATTTCAGCATGAATCAGGAATCTTGAGACCTTACTTTATGAGGAGTTGGAGAC 1380  
Db 1321 GAAGCTATTTCAGCATGAATCAGGAATCTTGAGACCTTACTTTATGAGGAGTTGGAGAC 1380  
QY 1854 acacgttgtatatattgaagaatcgaagacagacacataatacatctacccctcagga 1913  
1381 ACACGTTGTATTTATTTAAGATCAAGCAAGCAACCATATTAACATCTTACCTCCAGGA 1440  
QY 1914 atcacatgagtccgcctctgtatctcaagagagattcccaaaaggtgttaaacaattgaa 1973  
1441 ATCACATGATGCTCGTCTTTGTATTCAGGAGAGATTACCAAAAGGTGTAACATTTGAA 1500  
Db 1441 ATCACATGATGCTCGTCTTTGTATTCAGGAGAGATTACCAAAAGGTGTAACATTTGAA 1500  
QY 2034 gggccaactaaatcaaatcctcgtggtccgaccgctattactcgtattcgttaataag 2093  
1561 GGGCCAACATTAATCAATCCTCGTGCCTGACCGGCTATTATCTGTAGTTCTGTTAATATG 1620  
QY 2094 gagagagatcagcttcaagacatcattggccctcctcctcatctgcataaagaatctgta 2153



```
|||||
Db 1621 GAGAGATCTAGCTTAGGAGCTATTGGCCCTCTCCATCTGCTACAGAAATCTGTA 1680
OY 2154 gatcaagaagaacacgaataatgtcacagaagaagaaatgcaacgtgtttctgtatt 2213
Db 1681 GATCAAAAGAGAAACCAAGATATGTGCAGACANAGAGAAATGTCATCTGTTTTCTTATTT 1740
OY 2214 gatagaacccaagcttggtaacctcacagaagaatatacaagcttcttcccaaccagt 2273
Db 1741 GATGAGAACCAGACTGGTACTCTCAGAGATATACAAAGCTTCTCCCAATCAGCT 1800
OY 2274 gggatgagcttggagatccagaattccaagcttccaacatcattgacaagaatcattgac 2333
Db 1801 GGAGTGAAGCTTGAGATCCAGAGTCCAAAGCTTCCAAATCATATGCAAGCATCAATGGC 1860
OY 2334 tatgttttgaatgttgcagtgttgcagtgttgcagtgttgcagtgttgcagtgttgcagt 2393
Db 1861 TATGTTTTGATTAATTTGCAATGTGTGCAATGTGTGCAATGTGTGCAATGTGTGCAAT 1920
OY 2394 ctaagcatgtgagacagactgacttcttctgttcttcttcttcttcttcttcttcttct 2453
Db 1921 CTAAGCATTTGGAGCAGACACTGACTTCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTT 1980
OY 2454 cacaaaaatgtctctatgaagacacactcaacctatccattctcaggagaaatgttcttc 2513
Db 1981 CACAAAATGGTCTATGAAGACACACTCACCCTATTCCTCATCTCAGGAGAAATGCTCTTC 2040
OY 2514 atgtcgaatggaacccaagcttctatgatttcttgggttgcacaaactgagcttctggagac 2573
Db 2041 ATGTGATGGAAACCCAGGCTATGATCTGGGGTCCCAACTCAGACTTCTGGAGAC 2100
OY 2574 agagacatgacccgcttactgaagtttctagttgttgaacaaagacatgagatattac 2633
Db 2101 AGAGGATGAGCCGCTTACTGAAGTTTCTAGTTGTGACAAAGAACACGTGATTTATTTAC 2160
OY 2634 gaggaacgttatgaagatatttcagcacttctgtgtgaaacaaatgacatgaaacca 2693
Db 2161 GAGGACAGTTATGAAGATATTTCAGCATCTTGCTGATAAACAAATGCCATTGAAACCA 2220
OY 2694 agaaagcttccccagaatt----- 2711
Db 2221 AGAAGCTTCTCCAGGATCCTTCTTGCTTGGGATACCACTATGCTACTCAGATACCAAAA 2280
OY 2712 ----- 2711
Db 2281 GAAGAGTGAATCCCAAGAAAGTACACAGAAAAACAGCTTTTAAAGAAAAGGATACC 2340
OY 2712 ----- 2711
Db 2341 ATTTTGTCTCTGAACGCTTGTGAAGCAATCATGCAATAGCAGCAATTAATGAGGACAA 2400
OY 2712 -----tcaagacacctagcacttaggcaagaatattat 2747
Db 2401 AATAAGCCCAATAAGAAATCAGTCACTGGGCAAAAGCAAGSTAGGACTGAAAGGCTGTGCTCT 2460
OY 2748 gccaccccccagctcttgaacacgcatcaacggygaataactcgttactactcttcagtca 2807
Db 2461 CAAAACCCACAGTCTTGAAACGCCATTCACGCGGAATTAATCGTACTGCTTTCAGTCA 2520
OY 2808 gatcaagaagaatattgactatgataccaatcagttgaaatgaagaagaagaatttc 2867
Db 2521 GATCAAAAGAGAAATTGACTATGATGATACCATATCATGTTGAAATGAAAGAGAGATTCTT 2580
OY 2868 gacattatgatagagaatgaagaatcagagcccccgagcttccaagaagaacacgaagac 2927
Db 2581 GACATTTATATAGATGAGATGAAATACAGAGCCCGCAGACTTTCAAAAGAAACACGACAC 2640
OY 2928 tatttattgtcgaagtgaagagctcttgggataatgggaatgagatgagctcccaacatgt 2987
Db 2641 TATTTTATTTGCTGAGTGAAGAGCTCTGGGATTTATGGAGATGAGAGCTCCCAATGTT 2700
OY 2988 ctaagaagaaggtctcagatgtgagatgtccctcagttcaagaagaatgttttccaggaa 3047
|||||

|||||
Db 2701 CTAAAGAACAGGCTCAGAGTGGCAGTGTCCCTCAGTTCAAGAAAGTTGTTTTCCAGAA 2760
OY 3048 ttactgaatgcttcccttactaacgacctataccgttggagaactaaatgaacatttggga 3107
Db 2761 TTTTACTGATGGCTCTTTACTACAGCCCTTATACCTGGAGAACTAAATGAACATTTGGGA 2820
OY 3108 ctccctggggccaalatalaagaagcagaagtgtgaagataaatalcatgttacttcaagaat 3167
Db 2821 CTCCTGGGGCCATATATATGAAGCAGAAAGTTGAAGATATATCATGTAATCTTCAAAAT 2880
OY 3168 caagccttgccttcccttcttcttcttcttcttcttcttcttcttcttcttcttcttct 3227
Db 2881 CAGGCTCTCGCTCTATCTCTTATTTCTAGCCCTTATTTCTTATTTCTTATTTCTTATTTCT 2940
OY 3228 caagaagcagaacccatagaanaaaacttgcagagccatgaagccatgaacccaacttacttgc 3287
Db 2941 CAAGGAGCAGAAACCTAGAAAAAATCTTGTCAAGCTTATGAACCAAACTTACTTTTGG 3000
OY 3288 aaagtgcacatcatalatgagcccaataagaatgagttgactgcaagaccttggcttat 3347
Db 3001 AAAGTCAACATCATATGAGCACCCACTAAAGATGAGTTTGACTCAAAAGCCTGGGCTTAT 3060
OY 3348 ttctctgttgttgcacttggaaaaaagatgtgcactcaggtccgtatgttgaccccttctgtc 3407
Db 3061 TTTCTGTATGTTGACCTTGAAAAAGATGTGCATCAGGCTGATTTGACCCCTTCTGTC 3120
OY 3408 tgcacacttaacacactgaacccctgcataatggagagaagaatgacagatgagaaattgct 3467
Db 3121 TGCCACACTAACACACTATACCTTGTCTATGGAGACAGTGCAGATGAGAAATTTGCT 3180
OY 3468 ctgtttt---tcacacatcttgatgagacccaagaactgtgttacttcaactgaanaatltgaa 3524
Db 3181 CTGTTTTTCTCACACATCTTTGATGAGACCAAAAGCTGTGCTTCTCACTGAAATATGAA 3240
OY 3525 agaaactcgaaggtccctcgtcaatatccagaatggaatacccaactttaaagaagaattat 3584
Db 3241 AGAACTCAGAGGCTCCCTGCAATATCCAGATGGAAGATCCCATTTTAAAGAAATTA 3300
OY 3585 cgcttccatgcacatgaatgagctacataatgatacacttactgttgaatgagctcag 3644
Db 3301 CGCTTCCATGCATCAATGAGCTATGATATGATGATGATGATGATGATGATGATGATGATG 3360
OY 3645 gatcaaaagattcgaatgttcttgcagcatgagcaatggcaatgaanaacatccatctcat 3704
Db 3361 GATCAAAAGATTCGATGTATCTGCTCAGATGGGAGCAATGAACATCCATTTCTATT 3420
OY 3705 catctcagtgagacatgttctcagctgtacgaaaaaagaagaatgaataatggaactgtac 3764
Db 3421 CATTTCACTGACATGtgttctcagctgtacgaaaaaagaagaatgaataatggaactgtac 3480
OY 3765 aatctctatccagatgttcttgcagcagtggaatgttcatccatccaaagcttgaatttgc 3824
Db 3481 AATCTCTATCCAGGCTTTTGAAGACAGTGAATGTTTACCATCAACCTGGAAATTTGG 3540
OY 3825 cgggttggaatgccttatgtgcagacatctacatgccttgcagcagcatttcttgcgtg 3884
Db 3541 CGGGGGAATGCTTATTTGCGGACATCTACATCTGCGGATGAGACACATTTTCTGCTG 3600
OY 3885 taacgcaataaagtgtcagagctcccttgcagatgtgcttctgcagacatcgaagaatttgcag 3944
Db 3601 TACAGCAATTAAGTGTCAACATCCCTCGGAATGCTTCTGACACATTAAGATTTTTCAG 3660
OY 3945 attacagcttcaagacaatalatgagcagtgaggcccccaagcttgcagcagcatttcttgc 4004
Db 3661 ATTACAGCTTGAAGACATATATGACATGAGGCCCCCAAAAGCTGGCCAGACTTCATTAATTC 3720
OY 4005 ggaatcaatcaatgtccttgcagacccaagagagccttcttctgcagcagcagtgatgtgtgt 4064
Db 3721 GGATCAATCAATGCTGAGACCAAGAGGCTTTCTTGATGATCAAGTGATGCTGTTG 3780
OY 4065 gcaccaatgatattatcagcagcacaagaaggtgtccgttcaagaatgttccagcctc 4124
Db 3781 GCACCAATGATTTATTCAGCGGATCAAGACCCAGGCTCCGCTCAGAGATTTCTCAGGCTTC 3840
|||||
```



QY 4125 tacatctcagttatcatcatatagttctttagtggaagatggaactatcga 4184  
|||||  
Db 3841 TACATCTCTCATGTTATCATATATATATAGTCTTATGGGAAGATGGACACTTATCGA 3900  
QY 4185 ggaattccacttggaaccttaatgtctctcttcttgcaatgtaattcctcggataaa 4244  
|||||  
Db 3901 GGAATTTCCACTGGACCTTATATGCTCTCTTTGGCAATGGATTCATCTGGGATATAA 3960  
QY 4245 cacaatatttcaacctccaattattgtctcgatataccglttgaccaactaatat 4304  
|||||  
Db 3961 CACATATTTTTTAACCCCTCCATATTTATGCTGATACATCCCTTTGACCCCAACCATATAT 4020  
QY 4305 agcatctgcaacctctcttgcaatgagttgagtggtggtggtttaaataagttgagcag 4364  
|||||  
Db 4021 AGCATGTGCACGACTCTCTTCGATGAGTGTATGGGCTGTGTATTAATAGTTGACGCTG 4080  
QY 4365 ccattgggaatggaagtaaaagcaatatacagaatgcacagatactgcttcaactt 4424  
|||||  
Db 4081 CCATTTGGGAATGAGAGTAAGTAAGCAATATACATGACAGATTACTGCTCATCTACTTT 4140  
QY 4425 accaatatgttgcacctggtctccttcaaaagctcgacttcaactccaaggagag 4484  
|||||  
Db 4141 ACCAATATGTTTGGCCACTGTCTCTTCAAAAGCTGACTTCACTTCAAGGAGGAGT 4200  
QY 4485 aatgcctgggaacctctgagtggaataatccaaaagatggtgctgcaagtggaactcagaag 4544  
|||||  
Db 4201 AATCCCTGGAGACCTCGAGTGAATTAATCAAAAGATGGCTGCAAGTGGACTTCCAGAG 4260  
QY 4545 acaatgaagatcacaggaagtaactactcaaggagtaaatctctgtctacacagatg 4604  
|||||  
Db 4261 ACATGAAGATCACAGAGTAACACTACAGGAGTAATAATCTGCTTACAGAGATGAT 4320  
QY 4605 gtgaagagttcctcatctccagcagatcaagaatgagcatgagtaactcttcttcaag 4664  
|||||  
Db 4321 GTGAGAGAGTCTCATCTCAGCAGTCAAGATGGCATGAGTGCATCTCTTTTTCAG 4380  
QY 4665 aatggcaagtaagaagtttttcagggaataatcaaacctctcacacttggtgaactct 4724  
|||||  
Db 4381 AATGGCAAGTAAGAGTTTTCAGGGAATCAAGACTCTTCAACCTGTGTGAACCT 4440  
QY 4725 ctgaaccacacgttactgactcgtactccttgaattcaaccacagattggtgacag 4784  
|||||  
Db 4441 CTAGACCCACCGTACAGACTCGCTACTCTTGATTCACCCCGAGAGTGGGTGCACAG 4500  
QY 4785 atgcccctgagatgagagttctggtgctgcagagcagaagcctctac 4832  
|||||  
Db 4501 ATTGCCTGAGAGATGAGGTTCTGTGGCTGCGAGGACAGACGACTCTTAC 4548  
RESULT 6  
LOCUS 108641 4551 bp PAT 02-DEC-1994  
DEFINITION Sequence 1 from Patent WO 8800831.  
ACCESSION 108641  
VERSION 108641.1 GI:588647  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 4551)  
AUTHORS Pasek,M.P.  
TITLE DNA SEQUENCES CODING FOR MODIFIED FACTOR VIII:C AND MODIFIED FACTOR  
VIII:C-LIKE POLYPEPTIDES AND PROCESSES FOR PRODUCING THESE  
POLYPEPTIDES IN HIGH YIELDS  
Patent: WO 8800831-A 1 11-FEB-1988;  
JOURNAL Location/Qualifiers  
FEATURES  
source 1..4551  
BASE COUNT 1345 a 1003 c 1004 g 1198 t 1 others  
ORIGIN

Query Match 51.4%; Score 4085; DB 10; Length 4551;  
Best Local Similarity 95.1%; Pred. No. 0;  
Matches 4325; Conservative 0; Mismatches 31; Indels 192; Gaps 3;  
QY 477 gccacccgaagaatactactggtgagtgagtggaactgtaactatggaactatgcaagatg 536  
|||||  
Db 4 GCCACGAGAAATATCTACTGSGTGCACTGGAGACTGCTATGAGCAATATGCAAAAGTAT 63  
QY 537 ctggtgagctgctctgtygagcgaagattccctccatagatgcaaaaactcttccatc 596  
|||||  
Db 64 CTCGGTGAAGCTGCTGTGGAGCAAGATTTCTCTTAAGTGCAGAAATCTTTTCCATTC 123  
QY 597 aacacctcagtcggtgtaaaaaaagactglttltagaatccaggttacccttcaac 656  
|||||  
Db 124 AACACCTTAGTCTGTATCAAAAAGACTTGTGTTAGATTCACAGGATCACTTTTCAAC 183  
QY 657 atcgttaagccaaagccacccttgatggtgtctgtcaggttccatccatccagctgaagt 716  
|||||  
Db 184 ATCGCTAAGCCAGGCCACCTGGATGGGTCTGCTAGGTCCTACATCCAGGCTGAGGTT 243  
QY 717 tatgatacagtggtcaatatacacttaagaacatggtcccatccctgtaagttcaagct 776  
|||||  
Db 244 TATGATACAGTGTGCTATTAACACTTAAGAACATGGCTTCCATCCGTGACGCTTCATGCT 303  
QY 777 gttggtatcctactgtaaaagcttctgaggaagctgtaatatgatatacagaacagtaaa 836  
|||||  
Db 304 GTTGGTGTATCTTACTGGAAGCTTCTGAGGAGCTGATATGATGATCAACCAAGTCAA 363  
QY 837 agggagaagaagaatgataaagcttccctgttgaagccaataatgltctggaagctc 896  
|||||  
Db 364 AGGAGAAAGAAAGATGAATGAAGTCTTCCCTGTGGAACCATACATATGCTGGAGGTC 423  
QY 897 ctgaagaagaatggttccaatgagctctgaagcactggtccttaccatcaatcttct 956  
|||||  
Db 424 CTGAAGAGAAATGTCCATATGCTGACCCACTGTGCTTACTTACTATCTTACTTCT 483  
QY 957 catgtgaacctgttaaaagacttaattcaagcctcaattgagccctactgattgta 1016  
|||||  
Db 484 CATGTGACCTGTGAAGACTTAATTCAGGCTCATTTGAGGCTTACTGATGTAAGA 543  
QY 1017 gaaagagctctgccaagaaagaaacacagacacttgcacaaattataactcttctg 1076  
|||||  
Db 544 GAAGGAGTGTGGCCAGGAAGAAACACACACTTGCACAAATTTATPACTTTTCT 603  
QY 1077 gtaattgataagggaaaagcttggcaactcagaacaaagaactccctgaatgcagatagg 1136  
|||||  
Db 604 GTATTTGATGAAGGAAAGTTGGCACCTCAGAAACAAAGAACTCTTGATGACAGATAGG 663  
QY 1137 gatgtgcatctgcttgagggccttgacctaaatgacacagatcaatggtatgtatgaaacag 1196  
|||||  
Db 664 GATCTGCATCTGCTCGGGCTTGCCCTTAATATGCAACAGTCAATGGTTATGTAAACAGG 723  
QY 1197 tctctg---ccaagttctgattgtgacagacagaagaatcagttcttggcatgtatgta 1253  
|||||  
Db 724 TCTCTGTACACAGGTGTGATGTGATGCAACAGAAATCAGTCTATTTGGCATGTGATGGA 783  
QY 1254 atgggacacactccttgaaagtgacactcaatatctctgaaagtgacacaaattctgtgag 1313  
|||||  
Db 784 ATGGGCAACACTCTCTGAAGTGCACCTCAATATTTCTGCAAGGTCAACATTTCTGTGAGG 843  
QY 1314 aaccatgcgaagtgctcttggaatctgcgaataacttccctactgctcacaagctc 1373  
|||||  
Db 844 AACCATGCGACGGGCTCTTGGAATCTGCGCAATTAATTTCTTACGCTCAACACATC 903  
QY 1374 ttgatgacacttgagagtttctactgttltgtlataatcttccaccacaatgagtg 1433  
|||||  
Db 904 TTGATGACCTTTGAGAGTTCATCTGTTTGTCAATATCTTCCACCAACATGATGGC 963  
QY 1434 atggaagcttatgtcaaatgatacagctgttccagaagaaacccaactaagaaatgaaat 1493  
|||||  
Db 964 ATGGAAGCTTATGTCAAAAGTATGACAGCTGTCCAGAGAACCCCAACTACGATGAATAAAT 1023  
QY 1494 aatgaaagaaggaagactatgatgatcttactgattcttgaaatgagtgagtgcaag 1553



```
|||||
Db 1024 AATGAAGAGGGAAGACTATGATGATCTACTGATTCGAATATGATGTGGCAG 1083
Oy 1554 ttgatgatgaacaatcccttccttccaaatcgcgtacgttgcgaagaagacac 1613
Db 1084 TTTGATGATGACACTCTCCCTCTCTTATCCAAATTCGCTCAGAGAGCAATCCT 1143
Oy 1614 aaacttggagcatatacatgtcgtcgaagaagaagactggagactatgctccctagtc 1673
Db 1144 AAAACTTGGATACATTAATTCCTGCTGAGAGAGGAGACTGGAGCTATGCTCCCTAGTC 1203
Oy 1674 ctgcgcccgatgacagaagtataaaagtcataatltgaaacaatggccctcaaggat 1733
Db 1204 CTCGCCCCCGATGACAGAGATTATAAAGTCAATATTGGAACAATGGCCCTCAGGGATT 1263
Oy 1734 ggttgggaagtacaanaagtcgactttagtgaacacagaatgaacacttgaagctgt 1793
Db 1264 GGTGAGAGATCAAAAAGTCCGATTTATGGCATACACGATGAACCTTTAGACTCCT 1323
Oy 1794 gaagcatatcagcatgaatcaggaatccttggacccttacttacttctggggagctggaag 1853
Db 1324 GAACCTATTCAGCATGATGATCAGGAATCTTGGGACCTTACTTTATGGGGAAGTTGGAGAC 1383
Oy 1854 acaactgttgaattatattaaagaatcaagcaagcagaacataatacctaccctcaagga 1913
Db 1384 ACACGTGTGATTATATTTAAGAAATCAGCAAGCAGCATATACATCTACCCCTCAGCGA 1443
Oy 1914 atcactatgtccgctcccttctgtatcacaaggaatcacaagaagtgttaaacatttgaag 1973
Db 1444 ATCACTGATGTCCGCTCTTTGATTTCAAGGAGATTAACAAAAGGTGTAATAACATTTGAAG 1503
Oy 1974 gatttcccaatctctgcagagaataatcacaataataatgagcagctgacttgaagaat 2033
Db 1504 GATTTTCCAAATTCGTGCCAGAGAAATATTCAAAATATAATGACAGTGACTGTAGAAAT 1563
Oy 2034 gggcccaactaaatcaagatccctcggtgcctgacccgctattacttgaagttcgtaatatg 2093
Db 1564 GGGGCACTAAATCAGATCCCTGGGCTCGAGCCGCTATTAATCTTAAGTTTAAATAG 1623
Oy 2094 gagaagatcagcttcaaggaactatggccctcccaatcgtgtacaaagaactgtga 2153
Db 1624 GAGGAGATCTTAGCTTCAGGACTCATTTGGCCCTCTCCATATGCTACAAAGAACTGTGA 1683
Oy 2154 gatcaagaagaacaacagataatgtcagacaagaagaatgtlcatcctgttcttctglat 2213
Db 1684 GATCAAGAGAGAAACAGATATATGCACAAAGAGAAATGTCATCCTGTTTCTGATATT 1743
Oy 2214 gatgaagaaccgaagctgtgtacctcaagaagaataatacaagcttctccccaatccagct 2273
Db 1744 GATGAGAAACCAAGCTGTACTCTCAGAGAAATATACAAAGCTTCTCCCAATCCAGCT 1803
Oy 2274 ggaatgcagcttggagatccagaagtltccaagcctccaacatcatgacagcatcaatgac 2333
Db 1804 GGAATGCAGCTTGAAGATCCAGAGTTCCAAAGCTCCAAATCATATGCACAGCATCAATGGC 1863
Oy 2334 tatgttcttgaatgttgcagltgtcagltgttgcagltgaggtgggcatctactgtgacat 2393
Db 1864 TATTTTGTGATGATTTGCAATTTGCAATTTGTTGCAATGAGTGGCATACCTGTAAT 1923
Oy 2394 ctaagaatltgagacagacagacttcccttctgtcttcttcttctcggatataccctcaaa 2453
Db 1924 CTAAGCATTTGGAGACACAGACTCTCTTCTGTCTTCTCTCTGGAATATACCTTCAAA 1983
Oy 2454 cacaaaaatgtctctatgaagacacactccctatcccatlctgaaggaagactgtcttc 2513
Db 1984 CACAAATGGCTATGAAGACACACTCACCTATTTCCATTTCTCAGGAGAAACTGTCTTC 2043
Oy 2514 atgtcgaatgaaaaacccaggtctatggaatccttgggggtgcacaactcagacttgcgaac 2573
Db 2044 ATGTCGATGGAACCCAGGTCTATGATGATCTGGGGTGCCAACTCAGACTTTTGGAAAC 2103
Oy 2574 agagagcatgaaccgcttactctgaaggttctcagttgtgacaagaacaactcgtgatactaac 2633
|||||

Db 2104 AGAGCAGTACCGCCTTACTGAAGGTTTCTAGTTGTGACAAAGACACTGTGATTTATAC 2163
Oy 2634 gagaagcatatgaagataatcacaacatctgtcgaagtaaaaaacatgcatcgaacca 2693
Db 2164 GAGGACAGTTATGAAGATATTTTACACATATCTTGCTGAGTAAAAACAATGCCATTGMAACA 2223
Oy 2694 agaagcttctccagaat----- 2711
Db 2224 ACAAGCTTCTCCAGAGATCCTCTGCTTGATTAACCACTATGTGACTCAGATACCAAAA 2283
Oy 2712 ----- 2711
Db 2284 GAAGAGTGAATCCCAAGAGAAATGTCACCAGAAAAAACAGCTTTTAAAGAAAAAGATACC 2343
Oy 2712 ----- 2711
Db 2344 ATTTGTTCCTTAAGCTTTGTGAAGCAATCATGCAATAGCAGCAATTAATGAGGACAA 2403
Oy 2712 -----tcaagacaacctagactagcgaaagaatltat 2747
Db 2404 AATTAAGCCCGAAATAGAAAGTCACTGGGCAAAAGCAAGGTAGACTGAAGAGCTGTGCTCT 2463
Oy 2748 gccaccccaaccgctcttgaagcgcatcaaaoggggaataactcgttactactcttcagta 2807
Db 2464 CAAAACCCACAGCTTTGAACGCCATCAGCGGAATTAACCTGTACTACTCTTCAGTCA 2523
Oy 2808 gatcaagaagaattgactatgatataccatcatcagtttgaatgaaagaagaattt 2867
Db 2524 GATCAAGAGAAATTTGACTATGATGATACATATCAGTTGTAATGMAAGAAAGATTTT 2583
Oy 2868 gacattatgatgatgaatgaatgaatgaagcccccgacgttccaagaagaacaagaac 2927
Db 2584 GACATTTATGATGAGAGATGAAGAAATCAGAGCCCGCCAGCTTTCAAAAGAAACACGACAC 2643
Oy 2928 taattatgtcgaagtgagagagctcgtggattaatggatgagtagtcccccaatglt 2987
Db 2644 TATTTTATTTGCTGAGTGAGAGGCTGTGGATTATGGAATGAGTAGTCCCAATGTT 2703
Oy 2988 ctaagaacaagggctcaagatgagtgatccctcaagttcaagaagaagtgttctccagaa 3047
Db 2704 CTAAAGAACAGGGCTCAGAGTGGAGTGTCCCTGAGTTCAAGAAAGTTTGTCCAGGAA 2763
Oy 3048 ttactatgtgctccttactcaagccctataacogltgagaactaaatgaacatttggga 3107
Db 2764 TTTACTGATGGCTCCTTACTACAGCCCTTATACCGTGGAGAACTAAATGAACATTTGGGA 2823
Oy 3108 ctccgtggggccatatataagagcagaagltgaagaataatcatatgtaacttgaagaat 3167
Db 2824 CTCTGGGGGCCATATATATGAAGCAGAAATGAATATATCATATGATATCTTCAGAAAT 2883
Oy 3168 caggcctcgtccctatccttctatctagccttatttcttataatgaagaatcagaag 3227
Db 2884 CAGGCTCTCGTCCCTATATTCCTTATTTCTAGCCCTTATTTCTTATGAGGAAGATCAGAGG 2943
Oy 3228 caaggaagcaaacctagaaaaaactltgcaagcctaaatgaacccaaacttacttgg 3287
Db 2944 CAAGGAGCAGAACTTGAAGAAACTTTGTCAAGCTTAATGAACCAAACTTACTTTGG 3003
Oy 3288 aaagtcaaatcatatagaccccaactaaagaatgaatgttgccttcaaaagccttggactat 3347
Db 3004 AAAGTGCAACATCATATGAGCCCACTTAAGATGATGATGCTGCAAAAGCCTTGAGCTTAT 3063
Oy 3348 tctcttgatgttgaccttgaanaaagatgtgcaactcaagcctgacttgaacccctctgtc 3407
Db 3064 TTTCTGTGATGTGACTTGGAAGAAAGATGTGACATCAGGCTGATTTGGACCCCTTCTGTGC 3123
Oy 3408 tgcacactaacaacatgaacccctgtcactatgaggaagaagtgaacgtatcagaatltgct 3467
Db 3124 TGCCACACTTAACACACTGAACCTGCTCATGAGAGACAAAGAGATGACAGAAATTTGCT 3183
Oy 3468 ctgtttt---tcaacatcttgaatgagccaaagcctgttacttccacttgaanaaatatgaa 3524
Db 3184 CTGTTTTCCTCACCATCTTTGATGAGACCAAAAGCTGTGTCTTCACTGAGAAATATGGA 3243
|||||
```



QY 3525 agaacgtcagggctccctcgaatatccagatggaagatcccaattttaagaagaattat 3584  
|||||  
Db 3244 AGAAACGACAGGGCTCCCTGCAATATCCAGATGGAATCCCACTTTAAAGAAATAT 3303  
QY 3385 cgcctccatgcaatcaatggtctacataatgatactactggtctagtaatggtcag 3644  
|||||  
Db 3304 CGCTTCATGCAATCAATGAGCTGATGATGATGATGATGATGATGATGATGATGAT 3363  
QY 3645 galcaaaaggttgcgtatctgctcgaatggtcgaatgaaatccatctatt 3704  
|||||  
Db 3364 GATCAAAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3423  
QY 3705 cattcagtggaatggtctcactgtacgaaaaaagaagatataaatggtcgttac 3764  
|||||  
Db 3424 CATTTGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3483  
QY 3765 aatctctacaggtgttcttgagacagtggaatgttaccatccaaagcttggaattg 3824  
|||||  
Db 3484 AATCTATATCCAGGTGTTTTGAGACAGTGAATGTATTACATCCAAAGCTGGAAATTGG 3543  
QY 3825 cgggtggaatgcttatttgcgagacatctacatgctggtggaatgacacatttctgtg 3884  
|||||  
Db 3544 CGGCTGGAATGCTTATTTGGCAGACATCTACATGCTGGGATGAGCACATTTTCTGCTG 3603  
QY 3885 tacagcaataagtgctcagactcccccgtggaatggtctcgaacattagaattttcag 3944  
|||||  
Db 3604 TACAGCATATAGTGTGAGACTCCCTGGGAAATGCTTCTGAGACATTAAGATTTTCAG 3663  
QY 3945 attacagctcaggaacataatggaagctggtgccccaaagcttgccagactctatcc 4004  
|||||  
Db 3664 ATTAACAGCTTCAGAGCAATATGAGAGTGGGCCCAAAAGCTGGCCAGACTTCAATTATTC 3723  
QY 4005 ggtatcaatcaatgctcgtggaacacaaagagcccttcttgatcaaggtggtatcgttg 4064  
|||||  
Db 3724 GGATCATCATGCTCTGGAGACACCAAGAGACCTTTTCTTGATCAAGGTGATCTGTTG 3783  
QY 4065 gcaaccaatgattatcaacgacatcaagaaggtgccccgacagaagcttccacgctc 4124  
|||||  
Db 3784 GCACCATATGATTTATTCAGGACATCAAGACCAGGTCGCCGTGAGAAAGTTCTCCAGCTC 3843  
QY 4125 tacactctcagttatcatcatgatagtatctgtatggaagaagtggtcagactatcga 4184  
|||||  
Db 3844 TACATCTCTCATGTTTATCATCATGATGATGATGATGATGATGATGATGATGATGAT 3903  
QY 4185 ggaatctccactggaaccttaaatggtcttcttggaatggtggaatctatggtgataaa 4244  
|||||  
Db 3904 GGAATTTCCACTGGAACCTTAATGCTTCTTGCAATGTGATTCATCTGGGATAAAA 3963  
QY 4245 caaataattttaacctcaactatattgtctcgaatacgcgttgacacccaactatatt 4304  
|||||  
Db 3964 CAAATATTTTAACTCCCAATTAATTATGCTCGATACATCCGTTTGCAACCACTCATAT 4023  
QY 4305 agcatctgcagcaactctgcagatgagtgatggtggttgatttaaatggttgcaagatg 4364  
|||||  
Db 4024 AGCATTCGCGACACTCTTCGATGAGATTGATGGCGTGTATTAATAGTTCAGAGATG 4083  
QY 4365 ccattggaatggaatgaaagaataataatgataatgacagatctactctactctt 4424  
|||||  
Db 4084 CCAATGGGAATGGAGATGAAGCAATATCAAGATGACAGATTAATCTCTTATCTCTACTTT 4143  
QY 4425 accaataatgttgccacactggtctccttcaaaagctcgaaatcaactccaaggaagat 4484  
|||||  
Db 4144 ACCAATATGTTTCCACACTGCTCTCTTCAAAAGCTCGACTTCACTCCCAAGGAGAGACT 4203  
QY 4485 aatgcttgagacccctcagtgaaataatccaaaagtggtcgaatggaatggtcctcagaag 4544  
|||||  
Db 4204 AATGCTTGAGAGCTCTGAGGATTAATCCAAAGATGCTGCAAGTGAGACTTCCAGAAAG 4263  
QY 4545 acaatgaagaatcagaaggaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 4604  
|||||  
Db 4264 ACAATGAAGATCAAGAGATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAAT 4323

QY 4605 gtgaagagttccctcatctccacagagtcgaagatggtccatcagtgactctctttttcag 4664  
|||||  
Db 4324 GTGAAGAGATTCCATCTCCACAGCAGTCAAGATGGCCATGATGAGACTCTTTTTCAG 4383  
QY 4665 aatggaagaatgaaggttttttcaagggaatcaagaactctctcaacctgtgtgtaactct 4724  
|||||  
Db 4384 AATGGCAAGATGAAGGTTTTTTCAGGGAATCAAGACTCTTTCACACCTGTGTGAAGTCT 4443  
QY 4725 ctgagaccacccgttactcgtactcgtactcgaattcaaccccaagatgtgtgtgacag 4784  
|||||  
Db 4444 CTAGACCCACCCGTTTACGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4503  
QY 4785 attgcccctgagatggaaggtcttggtggtcgtggaagcaaggaacctctac 4832  
|||||  
Db 4504 ATTGCCCTGAGATGAGGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4551

RESULT 7  
AR034084  
LOCUS AR034084 5035 bp DNA PAT 29-SEP-1999  
DEFINITION Sequence 1 from patent US 5869292.  
ACCESSION AR034084  
VERSION AR034084.1 GI:5949689  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 5035)  
AUTHORS Voorberg,J.J.  
TITLE Hybrid proteins with modified activity  
JOURNAL Patent: US 5869292-A 1 09-FEB-1999;  
FEATURES  
source 1..5035  
BASE COUNT 1484 a 1127 c 1110 g 1314 t  
ORIGIN

Query Match 47.2%; Score 3748; DB 9; Length 5035;  
Best Local Similarity 87.4%; Pred. No. 0;  
Matches 4374; Conservative 0; Mismatches 60; Indels 570; Gaps 2;  
QY 415 ccaactgcaatagagctccacactgctctctctgtgctcttgcattgcattat 474  
|||||  
Db 30 CCAACATGGAATAAGACTCTCCACCTGCTTCTGTGCTCTTTGCGATTCTGCTTTA 89  
QY 475 gtgcacacgaagatactactggtgtgagtggaatgcaatggaatggaatggaatggaatg 534  
|||||  
Db 90 GTGCCACGAGAAGATTAATCTGCTGGTGCAGTGAAGTGCATGAGTGAATATATCAAAAGTG 149  
QY 535 atcttggtgagctgctgtgtggaagcaagatttccctcctagagtgcacaaatctttccat 594  
|||||  
Db 150 ATCTCGGTGAGCTGCTGTGAGCAAGATTTCTCTTACAGTGCACAAATCTTTTCCAT 209  
QY 595 tcaaacctcagtcgtgtacaaaagactgtgtgtgaaatcacaggttcacacttttca 654  
|||||  
Db 210 TCAACACCTCAGTGTGTACAAAAGACTGTGTTGATGATTCACAGGATCACCTTTTCA 269  
QY 655 acatgctaaagcgaagccacttgatggtgtctgtatggtctactacatccaggtgag 714  
|||||  
Db 270 ACATCCCTTAAGCCAAAGGCCACCTGTGATGGTCTGTGAGTCTTACCATCCAGGCTGAGG 329  
QY 715 ttatgatacagtggtcatctactaaagaacatggtctccactcctgtcaagcttcag 774  
|||||  
Db 330 TTTATGATACAGTGTGATTAATCACTTAAGAACATGAGCTTCCATCTGTGACGCTTCATG 389  
QY 775 ctgtgtgtatctactggaagctcttgaaggaagctgaatgataatgataatgataatgataatg 834  
|||||  
Db 390 CTGTGTGTATCTTACTGTGAAGCTTGTGAGGAGCTGAATATGATGATGATGATGATGATGAT 449  
QY 835 aaaggaagaagaatgataaagttctcccttgggtggaagccaatacatatgtcttgcaag 894  
|||||  
Db 450 AAAGGGAAGAAGAATGATTAAGTCTTCCCTGTGATGGAAGCCATATATATGTGTGACAGG 509



QY	895	lccTgaagagaaTgylcceaTggccctcTgaaccactTgTgcTtaaccatacctacTcTcTt	954
Db	510	TCCTGAAAGAAATGGCTCCAAATGGCCCTCTGACCACCTGTGCCCTTAACCTACTACATATCTTT	569
QY	955	cTcaTbtTgaacccTgTaaagaaactTgaatTcaacTccatTggagccctactaagTaTba	1014
Db	570	CTCATGTGACCTTGTRAAAGACTTGAAATTCAGGCCCTCATTTGGAGCCCTACTAGTaTbTa	629
QY	1015	gaagaaggagTcTgTccaaaggaaagaaacaaagactTgcaacaaTtiaTactactTtTg	1074
Db	630	GAGAAAGGAATCTGGCCAAAGAAAAGACACAGACTTGCACAAATTTACTACTTTTTCG	689
QY	1075	cTgTatTtTgaTgaagggaaaaagTtTgTcaTcTcaagaaacaaagaactcctTtTgcagTaTb	1134
Db	690	CTGTATTTTGTATGAAGGGAAGATTGGCTACCTCAGAAACAAAAACCTCTCTGTATGACGATa	749
QY	1135	ggagaTgctgaTcTgTccTcgggccTgagccTaaatTgacacaaTgaTgaTgTtatTbaaCa	1194
Db	750	GGGATCTGCAATTTGCTCGGGCCCTGGCCCTTAATATGCACACGTAAATGGTTATATGAACa	809
QY	1195	ggTcTcTgccaagTcTgTaTtTgTaTgTgccaaggaatcaTgTatTtTgTgaTtTgTa	1254
Db	810	GGTCTGCGCAGGTCTGTATTTGGATTCGACAGAGAAATCAGTATTTGGCATGTGATTGGAA	869
QY	1255	TgggcaaccactcctTgaagTgTcaTcTcaatTatctcTcgaagTgcacaatTctTtTgTa	1314
Db	870	TGGGCACACACTCTGTGAAGTCACCTCAATATTCCCGAAGGTCACACATTTCTTGTAAGa	929
QY	1315	accatTgcCaagTgTcTcTtTggaatTctcTgcCaatTaaactTtccTctTactTgTcTaaacactT	1374
Db	930	ACCATGTGCCAGGGCTCTTGGAAATCTCGCCAAATACTTTCTTACTGTCTAAACACTCT	989
QY	1375	TgaTgacctTtTgacagTtTcTactgTtTtTgTcaTatcTcTcTccacaaacTgaTgTgCa	1434
Db	990	TGATGACCTTGGACAGATTCTACTGTGTTGTGCATATCTCTTCCACACACATGATGAGCa	1049
QY	1435	TggaaagctatTgTcaaaTgTgaacagTgcTgcTcgaagggaaccccaactcTgaTlgnaaaTa	1499
Db	1050	TGGAACTTATGTCAAAAGTGAAGACAGCTGCTCAGAGGAACCCCAACTCGAAATGAAAAATa	1109
QY	1495	atGaaagacTggaagactTgaTgaTgaTatcTtTactgTatTtTcTgaatTggaTgTgTcaagT	1554
Db	1110	ATGAAATAGCGGAAGACTATGATGATGATATCTTACTGATTCTGAAATGGAATGTGGCAGCT	1169
QY	1555	TtgaTgaTgaCaacTcTcTcTcTcTtTcaTcaaaTtTgcTcaagTtTgccaagaagactcTca	1614
Db	1170	TGTATGATGACACTCTCTCTTATCCAAATTCCTCTAGTGGCCAAAGACATCTCTa	1229
QY	1615	aaactTggTgTacaTtTcaatTgTgcTgTgTgaagagggagTgTgTgTgTcTctTtagTcc	1674
Db	1230	AAACTTGGGTACTTTCATTTGCTGCTGAAGAGGAGACTGGGACTATGCTCTCTTAGTCC	1289
QY	1675	TcgcccccTgaTgCaagaTgTataaaagTcaatTtTtTgaacaaTgTgcTcTcaagTgaTgTg	1734
Db	1290	TGCGCCCCGATGACAGAGTATTAAGTCAATTTTGAACAATGGCCCTCAGCGGAATG	1349
QY	1735	gTgaagaagTacaaaaaagTccgaTtiaTtTgTcaTcacagaTbnaaacTtTaaagTcTgTg	1794
Db	1350	GTAGGAAGTACAAAAAAGTCCGATTTATGGCATTCACAGATGAACCTTTAAAGCTCGTG	1409
QY	1795	aaagTatTcaagTgaTgaTgaagaaTcTtTtTggaacTtTactTtTatTtTgggaatTtTggaCa	1854
Db	1410	AAAGCTATTCGCATGGAATTCGGAATCTTGGACCTTACTTTATTTGGGAACTTGGAGCa	1469
QY	1855	caactTgTatTatTtTaaagatTcaagTcaagTcaagTcaTaaacatTcaactcTcaagTga	1914
Db	1470	CACGTGTTGATTATTTAAATTCAAACGACAGACCATATTAACATCTAACCTCAGGGAA	1529
QY	1915	tcaactTgaTgcTgTcTctTtTatTtcaagagTatccaTaaagTgTtaaacaactTtTgaagT	1974
Db	1530	TCAGTGATGCTCCTCTTGTATTTCAAGGAGATTTACCAAAAAGGTGTAAACATTTGGAAG	1589

QY	1975	atttttcaattctcttcocggggaaatlctcaatlataaalggagcagtgtgaaagatg	2034
Db	1590	atttttccaaatttctcccgaggaaattattcAAATATAAATGAAGAGACTGTAGAATG	1649
QY	2035	ggccaacataaactgaatctctcgtgtgccttgaccgcgctattactagtttcgttaatg	2094
Db	1650	ggcccaattAAATTCAGATTCCTGGGCTGACCCTGATTAATTAATTAATG	1709
QY	2095	agagagatctagcttcagagactatctggccctctctcaatctcgtctcaaaagatctgag	2154
Db	1710	AGAGAGATCTAGCTTCAGGACTCATTTGGCCCTCTCCATCTGCTAACAAAGAAATGTAG	1769
QY	2155	atccaaagggaacacaataatgtccagaagaagaaatgtcaatccgttttcgtattg	2214
Db	1770	ATCAAAAGGAAACACAGATTAATGTCAACACAAGAGAAATGTCAATCCGTTTCTGTATTTG	1829
QY	2215	atgaaacccgaagcttggtaacctcaagagaaatatacaagcttctctcccaatccagctg	2274
Db	1830	ATGAAACCGGAAGCTGGTACTCTCAAGAAATATACAGCGTTTCTCCCAATCCAGCTG	1889
QY	2275	gagtgacgcttggaatctcgaagttccgaacccccaacccaactctgacagatcaatggct	2334
Db	1890	GAGTCCACTTGAAGATCTCAAGTTCCAACCCCTCAACATCAATGACACACATCAATGGCT	1949
QY	2335	atgttttgaatagtttcgaattgttgcagtttggtttgaataggtgtgcaatctgtatacttc	2394
Db	1950	ATGTTTGTGATAGTTTGCAGTGTCAATTTGTTTGGATGAGTGGCATGATCTGTACATTC	2009
QY	2395	taagcatctggagcaacagatgaactctctctgtctctctctctcgtataatccttcaac	2454
Db	2010	TAAAGATTGGAGCAACACATGACTCTCTCTCTCTCTCTCTGTGATTTACCTTCAAC	2069
QY	2455	acaaatgtctataagaagaacataccctatctccatctccaaagaaactgtcttca	2514
Db	2070	ACAAATGTGTTATGAAGACACACTCACTCACTTATTCATTTCTCAGAGAAACTGCTTCA	2129
QY	2515	tgctgaatgaaaccacgaatctatgtatctctgggtgtgcacaactcgaactctcggaa	2574
Db	2130	TGTCATGTGAACAAACCCAGGCTATGATGATTTGGGGGCCAACATCAGACTTTCGAGACA	2189
QY	2575	gaggaatgacgcctctactaagaattctcagttctgtgtgtaagaagaactgtgtattg	2634
Db	2190	GAGGATGTACCGCTTACTGTGAAGTTTCTTAGTTGTATTCAGAGGGGGGAGGACGACG	2249
QY	2635	aggaacagttatgaagatatctca-----gcaactgtctgagtataaaca	2679
Db	2250	ACTATCTGGACCTGAGAGATATTTAGTGAAGACGACGATACATCGACATCGCGACA	2309
QY	2680	atgcatctgaaaccaagaagcttctccagaatctcaagaacacctgacatgagcaaaagc	2739
Db	2310	GTCGATTTGAACCAAGAAAGCTTCTCCCAATTCAGACACCCCTAGCACTAGCAAAAAGC	2369
QY	2740	aatttaatgcaac-----	2752
Db	2370	AATTTAATGCCACCAACAAATTCAGAAATGACATAGAGAAGACTGACCTTGGTTTGGAC	2429
QY	2753	-----	2752
Db	2430	ACAGAACACTATGCTTAATAATCAAAATGTCTCTCTAGTATTTGTGATGCTTGTGC	2489
QY	2753	-----	2752
Db	2490	GACAGAGTCTACTCCACATGGGCTATCTTATCTGATCTCCAGAGACCAATATGAGA	2549
QY	2753	-----	2752
Db	2550	CTTTTTCGATGATCATCAACCTGGAGCAATAGACAGTAATAACGCGTGTGAATGA	2609
QY	2753	-----	2752
Db	2610	CACACTTCAGGCCACAGCTCCATCAGAGTGGGAGCATGTATTTACCCCTGAGTAGGCC	2669
QY	2753	-----	2752



Db	2670	TCCAATTACATTAAATGACAAACCTGGGGACAACTGCACATCCTCTTGGTGGATTAAC	2722
Qy	2753	-----	2752
Db	2730	ACTATGACTCAGATATCCAAAAGAGTGGAAATCCCAAGAGAATCACCAGAAAAA	2789
Qy	2753	-----	2752
Db	2790	CAGCTTTAAGAAAAAGATACATTTTGTCCCTGAACCTGTGAAAGCATTATGCA	2849
Qy	2753	-----	2752
Db	2850	TAGACGCAATTAATGAGGGACAAATATACCCGAATATAAGTCACCTGGCAAAAGCAG	2909
Qy	2753	-----	2784
Db	2910	GTAGACGTGAAAAGCGCTGCTCTCAAAACCCACCACTTTGAAAGCGCATCAACGGAAA	2969
Qy	2785	taactgtactactcttcagtcagatcaagaggaattgactatgatataccatataag	2844
Db	2970	TAACTCGTACTACTCTTCATCTCAGATCAAGAGAAATTCATATGATATACCATATAG	3029
Qy	2845	ttgaaatgaagaagaagatttgcacattatgatgatgaatgaaatcaagccccgca	2904
Db	3030	TTGAAATGAGAAAGACGAATTTTGACATTATGATGAGATGATAAATCAGACCCCGCA	3089
Qy	2905	gcttccaagaagaacacgacactatttatgtcagctgagagagctctggaattg	2964
Db	3090	GCTTTCAAAAGAAAACACGACACTATTTTATTTGCTCAGTGGAGAGCGCTTGAGATTATG	3149
Qy	2965	ggatgagtagctcccaatglttctaagaacaaggtctagaatggcagtgctccctaagt	3024
Db	3150	GGATGATATACCTCCCAACATGTTCTAAGAAACAGGGCTCAGAGTGCGAGTGCCCTCAGT	3209
Qy	3025	tcaagaagatgltttcccaagaaatttaacgatgagccttctaagccctataacg	3084
Db	3210	TCAAGAAAGTGTTTTCCAGAGATTTACGTAGTGTGCTCTTATCTCAGGCCCTTATACCTG	3269
Qy	3085	gagaactaaatgaaacatttgggaactccctggggccabatalaagacagaagtgaagata	3144
Db	3270	GAGAACTAAATGAAACATTTTGGGACTCCTGGGGCCATATATAGACACAAAGTTGAAGATA	3329
Qy	3145	atacatggraaacttccaagaatcagaagcctctgcgtccatctccctatctcagccta	3204
Db	3330	ATATCATGTGTAACCTTTCAGAAATCAAGGCCCTGTCGTCCATTCCTCTTATTCCTATGACCCTTA	3389
Qy	3205	tttcttcttgaggaagatcagaaggtcaagagacagaacctagaanaaaacttgcagaaccta	3264
Db	3390	TTTCTTATGAGGAATCATCAGAGCAGCAAGAGAACCTTGAAAAAATCTTGTCAAGCCTTA	3449
Qy	3265	atgaaaccaaaacttactcttcttggaaagtgcacaatcatatgtgcaccacataaagatgagt	3324
Db	3450	ATGAAACCAAAACTTACTTTTGGAAAGTCAACATCATATATGGCACCCACTAAAGTATGAGT	3509
Qy	3325	ttgacgtcaaaagcctggagctattctctgcagtgtgagaccttgaaaaaagatgtgcaactag	3384
Db	3510	TTGACTCAAAAGCCTGGGCTTATTTCTCTGATGTTGACCTGAAAAAATGTGCACTCAG	3569
Qy	3385	gacctatgtgacccctctgtgtctgcacaactaaacacatgaaacctgtcatgtggagac	3444
Db	3570	GCCTGATTTGGACCCCTTCTGTCGTGCGACACATTAACACATGAACCTGTCATGGAAGAC	3629
Qy	3445	aagtgaaagtaacaggaatttgcctctgttlltcaacactttgatagaagccaaagctgagt	3504
Db	3630	AAGTGAAGATGACGAGATTTTGGCTCTCTTTTTCACATCTTTTATATAGACCAAAAACCTGTG	3689
Qy	3505	acttcacttgaaaaatatgaaaaaagacctgaaagctccctgtgcaatatccaagatgagaatctc	3564
Db	3690	ACTTCACTGAAAAATATGGAAGAAATGTGAGGGCTCCCTGCAATATCCAGATGGAAGATC	3749
Qy	3565	ccacttttaaagaagatlatcgcttccatgcaatcaatgctacataatgataactaac	3624

Db	3750	CCACTTTTAAAGACATTAATTCGCTTCATGCATCAATAGGCTACACTAATTCGATACACTTAC	3809
Qy	3625	ctggtcttgtaatggctccaagataccaagatctcgaatgglatctgctcagcaatggcagca	3684
Db	3810	CTGGCTTAGTAATGGCTCAGATCAAAAGATTGCATGGATATCGCTCAGCATGGGCACCA	3869
Qy	3685	atgaaacacatcatctatctatcttcagatggacatggttctacgttaccgaaaaaaagg	3744
Db	3870	ATGAACAACATTCATCTCTATTCAATTCAGTGGACAACTGTCTACTGTACGAAAAAAGAGG	3929
Qy	3745	agtataaaatggcactgtacaaatctctacccaggtgttttttgaagacatggaaaatggttac	3804
Db	3930	ACTATTAATAATGGCACTGTACATCTCTATCCAGGtggTTTTTGAGCACTGGAAAAGTTTAC	3989
Qy	3805	catccaagcttgaaattggcgggtgagaaatgccttatggcagagatctacatgcttgga	3864
Db	3990	CATCCAAAGCTGGAAATTTGGCGGGTGAATAGCTTATTGGCGAGCATCTACATGCTGGGA	4049
Qy	3865	tgaagcacacttllctggtgtacagcaataagtgctcagactccctcctggaaatgctctg	3924
Db	4050	TGAGACACACTTTTCTGGGTGTCACACAAATAAATGTCAACACTCCCTGGGAATAGGCTTCTG	4109
Qy	3925	gacacattagagaattttaaagtttaagctttagagcaaatatggaagatgggccccaaagc	3984
Db	4110	GACACATTAGAGATTTTTCAGATTACAGCTTCCAGCAATATGAGACATGGGCCCCAAAGC	4169
Qy	3985	tggcagacctcatattccggatacaatcagcttgtagccaagagacccttctct	4044
Db	4170	TGGCCAGACTTCATTATTTCGGATATCAATCAATGCTGTGAGCACCAAGAGCCCTTTCTT	4229
Qy	4045	ggatcaaggttgatctgtttgacccaatgatattatcaagcatalcaagaaccaaggttgccc	4104
Db	4230	GGATCAAGGTGATCTGTTGGCACCAATGATTATTACGGCATCAAGACCAGCGGTGCC	4289
Qy	4105	gtcagaagttctccaagcctctacatctctcagtttcaatcatgtaatgctttgatggga	4164
Db	4290	GTCACAAGTTTCCACGCTTACATCTCTCACTTATCATATGTAATGCTGTGATGGGA	4349
Qy	4165	agaagtggcagactatccagaagaaattccacatggaaacttaatggttctctttggcaatg	4224
Db	4350	AGAAGTGCAGACTTATGAGAGAAATTTCCACTGGAACTTAAATGCTTCTTTTGGCAATG	4409
Qy	4225	tggaatctatctgggataaaacaaatalltttaacccctccaattatgtctcgatacatcc	4284
Db	4410	TGGATTTCATCTGGATTAATAACACATATTTTTTAAACCTTCACATTAATGCTCGATACATCC	4469
Qy	4285	gtttgcacccaactaatatataagcatctgcagcactccttcgcatggatgtgatggcgtg	4344
Db	4470	GTTTGCACCCAACCTCATTTATATACATTTCCAGACCTTCCTCGATGGAGTTGATGGCTGTG	4529
Qy	4345	atttaataagtttgagcaatgcataatgggaatgggaagataaagcaatatccagatgcacga	4404
Db	4530	ATTTAATAATGTTGGACGATGCCATTGGGAATGGAGATGAAGCAATATTCAGATGCACAGA	4589
Qy	4405	ttactgcttcatcctactttaaacaatgatgtttgcacacttggcttctcttcaaaagctcgac	4464
Db	4590	TTACTGCTTCATCCTTACCTTAAACCAATATGTTGGCACCTGGTCTCTTTCAAAAGCTGCAC	4649
Qy	4465	ttcacctccaaggagagatlaatgccctggagacctcagatgtaataatccaanaagatggc	4524
Db	4650	TTTCACCTCCCAAGGGAGGAGTATAGCTCTGGAGACCTCCAGGTAAATATCCAAAGAGTGGC	4709
Qy	4525	tgcagatggacttccaagaagcaatggaangtccaagggtaactacttaaggaagttaaat	4584
Db	4710	TGCAGATGGACTTCACGAAGAAATGAAGTACACGGAGTAACCTACAGGAGATGAAT	4769
Qy	4585	ctctcttcaaccagatgatgtgaagaggttcctcatctccagcagtcacaagtggcatc	4644
Db	4770	CTCTGCTTACCAAGCATGTATGTGAAGGAGTTCCTATCTCCAGCAGCTCAAGATGGCCATC	4829
Qy	4645	agtggaatctccttlltccaagaatggcagaagttaaagttlltccaaggaaatcaagaatcctc	4704
Db	4830	AGTGCACTCTCTTTTTCAGATGGCAAGTAAAGTTTTCAGGGAATCAAGACTCTCT	4889



QY 4705 tcaacacgtgtgtagactcttagaaccacgltactgactgcgtactcgaatccacc 4764  
|||||  
Db 4690 TCACACCTGTGGTGAACCTCTTAGACCCACGCTACTGACTGCTACCTTCGAATTGACCC 4949  
QY 4765 cccgaagcttggtgtgacacagatctgcgcgaggaatgtgaggtctctggcctgcgagacag 4824  
|||||  
Db 4950 CCCAGAGTTGGGTGCACACAGATTGGCCCTGAGAGTGGAGGTCTTGCGCTCGAGGACAGG 5009  
QY 4825 accctactgactgcagcagctc 4848  
|||||  
Db 5010 ACCTCTACTGAGGGTGGCCACTGC 5033  
RESULT 8  
AR071306 AR071306 5035 bp DNA PAT 18-FEB-2000  
LOCUS Sequence 1 from patent US 5910481.  
DEFINITION AR071306  
ACCESSION AR071306  
VERSION AR071306.1 GI:7222194  
KEYWORDS  
SOURCE .  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 5035)  
AUTHORS Voorberg,J.J.  
TITLE Hybrid proteins with modified activity  
JOURNAL Patent: US 5910481-A 1 08-JUN-1999;  
FEATURES Location/Qualifiers  
source 1..5035 /organism="unknown"  
BASE COUNT 1484 a 1127 c 1110 g 1314 t  
ORIGIN

Query Match 47.2%; Score 3748; DB 9; Length 5035;  
Best Local Similarity 87.4%; Pred. No. 0;  
Matches 4374; Conservative 0; Mismatches 60; Indels 570; Gaps 2;

QY 415 ccaccatgcaaatagagctctccaccctgctctctctgtgctcttgcgactctgctta 474  
|||||  
Db 30 CCACCATGGAATATAGAGCTCTCCACCTGCTCTTCTGTGGCTTTTGGCATTTGCTTTA 89  
QY 475 gtgcacacagaagatactactcgtggtgtagtgagaaatgltcatggtgactatagcaaatg 534  
|||||  
Db 90 GTGCACACAGAAGATATACCTGGGTGCAGTGGAACTGCATGGAGACTATATGCAAAAGTG 149  
QY 535 atctcgtgtgagctgctgtgtgagcgaagaattctccctcctagagtgccaaaaattcttccat 594  
|||||  
Db 150 ATCTCGGTGAGCTGCTGCTGTGACGCAAGATTTCTCTTAGAGTGCMAAAATCTTTCCAT 209  
QY 595 tcaaacactcagctcgtgtgacaaaagaagactcgttctgtagaattcaggtgttacccttca 654  
|||||  
Db 210 TCACACCTCTGCTGCTGTACAAAAGACTCTGTTGTAGAAATTCACGGATTCACCTTTTCA 269  
QY 655 acatcgttaagcgaagccaccctgtagtgggtcgtgtagtctcaaccatccagagctgag 714  
|||||  
Db 270 ACATCGGTAAACCAAGCGCACCTTGATGGGTCTAGTCTCTAGTCTCATCCAGGCTGAGG 329  
QY 715 ttatagtacagtggtgtacttaacttaagaacatggtctcccaatcctgtcagcttcaatg 774  
|||||  
Db 330 TTTATGATACAGTGGTCTTACACTTAAGAACATGGCTTCCATCTTGCACTTTTCATG 389  
QY 775 ctgtgtgtgtactcactggaagaagctctgagggagctggaatatgatatcagacagctc 834  
|||||  
Db 390 CTGTGTGGTATCTCTACTGGAAGCTTCTGAGGGAGCTGAATATATATGATCAGACAGTGC 449  
QY 835 aaagggagaagaagatgataaagctctccctggtgtagaagccatacatatgctgtgcag 894  
|||||  
Db 450 AAAAGGAGAAAGAGATGATTAAGCTTCCCTGGTGAAGCATATATATGCTGTGCAAGG 509  
QY 895 tccctgaagaagaatggttcaatggtccttgagccacgtgtgcttaccatcatcttctt 954  
|||||

Db 510 TCCGTGAAGAGAAATGGTCCATATGCGCTTGACCCACAGTGGCTTACCTATATATCTTT 569  
QY 955 cctcatgtgacctgtgtlaaaagacttgaattcagcgctcatgtgagccctactagtatga 1014  
|||||  
Db 570 CTCATGTGACCTGTGTAAGAAAGACTTGAATTCAGGCGCTCATTTGAGCCCTACTAGATGA 629  
QY 1015 gagaaggagagcttggccaaggaagaacagaccttgcacaaatttactacttcttg 1074  
|||||  
Db 630 GAGAAAGGAGTGTGGCCCAAGGAAAGACACAGACCTTGACAAATTTATCTACTTTTGG 689  
QY 1075 ctgtatttgatgaagggaagaagtgtgcaatcgaagaacaaagaaccccttgatgaagata 1134  
|||||  
Db 690 CTGTATTTGATGAAGGAGAAAGTTGGCACTCAAGAAACAAAGACTCTTGTGATGAGAGATA 749  
QY 1135 ggaatgtcatctgtctcggtcgtgacctaaatgcacacagtcgaatgttatgaaca 1194  
|||||  
Db 750 GGGATGCTGATCTGCTGCGGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 809  
QY 1195 ggtctcgcgaagctgtgatttgatgagccaaagaaatcagctatgtagatgtgaa 1254  
|||||  
Db 810 GGTCTGCTGCCAGTCTGATTTGGATGCCACAGGAATCAGTCTATTTGGCATGTGATGGAA 869  
QY 1255 tgggacacactcctgaatgtgcaatcaatatctctctggaagtgcaacaaattctgtgaga 1314  
|||||  
Db 870 TGGGACACACTCTCTGAAGTGCACTCAATATTCCTGGAAGTGCACACATTTCTGTGAGGA 929  
QY 1315 accatcgcaagcgctcctctggaatctcgccaataactctctactgctcaaacactct 1374  
|||||  
Db 930 ACCATTCGCCAGGCGTCTTGGAAATCTGCGCCATATACCTTCTTACTGCTCAAACTCT 989  
QY 1375 tgaatgaccttggagacagttctactgttltgtcatatctctcccaacaatgagtga 1434  
|||||  
Db 990 TGATGACCTTGAGACAGATTCTTACTGTTTGTGCAATATCTTCCACACCAACATATATGGA 1049  
QY 1435 tggaaagcttatgtcaaatgataagcgtgtccgaagggaaccccaactcgtgaatgaataa 1494  
|||||  
Db 1050 TGGAACTTATGTCAAGATGACACGCTCTCCAGAGAAACCCCAACTCGAATGTGAATAAATA 1109  
QY 1495 atgaaagacggaagaactatgatagtatctactgactctcgaatgtgaltgtgtcagct 1554  
|||||  
Db 1110 ATGAAGAAGCGGAACACATATGATGATGATCTTACTGATTCGAATGTGATGTGTCAGGT 1169  
QY 1555 ttgatgtgaaacactcctcctcttccaaatcgcctcgaatgtgcaagaagatcccta 1614  
|||||  
Db 1170 TTGATGATGACAACTCTCTTCTTATTCAAATTCCTGCTGCTGCAAGAACATGCTTA 1229  
QY 1615 aaacttgggtacatcacatctgtcgtggaaggaggagctgtgagctatgcttccctatcc 1674  
|||||  
Db 1230 AAACTTGGGTACATTACATTTGCTGTGAAGAGAGGAGACTGGGACTATGCTCTTAGTCC 1289  
QY 1675 tgcgcccgatgacagaagttataaaagtcaaatlttgaacaatgagccctcagcgatg 1734  
|||||  
Db 1290 TCGCCCGGATGACGAAGTTATATAAAGTCAATTTTGAACAATAGGCCCTCAGGGAATG 1349  
QY 1735 gtaggaagtlacaaaaaagtlccgattatggtacatacacagatgaaacotttgaagctgtg 1794  
|||||  
Db 1350 GTAGGAAGTACAAAATAATCCGATTTATGCGATACACAGATGAACCTTTAAAGCTGTGTG 1409  
QY 1795 aagctatcagatgaatcaagaatccttggagaccttactttaggggaagtgtgagaaca 1854  
|||||  
Db 1410 AAGCTATTTAGCATGAATCAGGAATCTTGGGACCTTTTATGAGGGAATGTGGAAGCA 1469  
QY 1855 cactgtgatlatatlttaagaatcaagcaagcaacacataaatacctacacagaa 1914  
|||||  
Db 1470 CACTGTTGATATATATTTAAGAAATCAGCAAGCAAGCAATATATATCTACCTGACGGAA 1529  
QY 1915 tcaatgagtgccgtccttltgtlatlcoaggagatataccaaagtgltlaaacaatttgaag 1974  
|||||  
Db 1530 TCACGTATGCTCGTCTTGTATTCGAAGAGATTACCAAAAGGTGTAAACATTTGAAGG 1589  
QY 1975 atttccaaatctgcgcgaaggaagaatatcaatatgaatgtagacgtgactgttgaagatg 2034  
|||||  
Db 1590 ATTTTCAATTCTGCGCAGGAAGAAATATTTCAATATTAATGACAGTACTGTAGAAGATG 1649



QY	2035	ggccaaactaaatcagatcctccgblgcctcgaccgcgtattaccctagttcgtglaataatg	2094
Db	1650	GGCCAACTAAATTCAGATCCTCGGTGGCTGACCCGCTATTACTGATTCCGTTAATATGG	1709
QY	2095	agagaatctcagcttcaagactcattggccctctctctatctgcataaagaatctgag	2154
Db	1710	AGAGGATCTTAGCTTCAGGACTATTGGCCCTTCTCTCATCTGCTACAAAGAACTGTAG	1769
QY	2155	atcaaagaagaaacagataatgtcagacaagaagaaatgcaaccgtttcttcgtatg	2214
Db	1770	ATCAAGAGGAAACCAATATATGTCAGACAAAGAGAAATGATCCGTGTTCTGTATTTG	1829
QY	2215	atggaagaacgaagcttggttacctcaacagagaataataaagctttctcccaatccagctg	2274
Db	1830	ATGGAACCGAAGCTGGTACTCTACAGAGAAATATACAGCTTTTCTCCCAATCCAGCTG	1889
QY	2275	gagtcagcttgaaagatccagagttccaaagctccacaacaaatgcaagaatcaatgct	2334
Db	1890	GAGTGAAGCTTGAGAGATCCAGAGTTCCAAAGCCCTCAACATCATGTGACAGCATCAATGCT	1949
QY	2335	atgtttctgatagtttgcagctgacgtctgctgttgcatagaagtgacatactgatactc	2394
Db	1950	ATGTTTGTGATGTTTGCATGCTTCAGTTGTTGGATGAGGTGCAATACGTAACTTC	2009
QY	2395	taaacattggaagcaacagactactcttcttgcctctctctctctgataactctcaac	2454
Db	2010	TAAAGATTGGAGCAACACAGACTCTCTTCTGTCTCTCTCTCTGTGATATACCTTCAAC	2069
QY	2455	acaaatggtctatgaaagacacactaccctatcccatctccagagaagaaactgctca	2514
Db	2070	ACAAAATGGTCTATGAAGACACACTCACCTCTATTCCATTCTCAGAGAAACTGCTTCA	2129
QY	2515	tgtcgaatggaaaacccaaggtcctatgattctcgggtgtgcacaactcgaacttcggaa	2574
Db	2130	TGTGATGAGAAACCCAGGCTATGTGATTTCTGGGGGCCACACACTCAACTTTGGAA	2189
QY	2575	gaagcagacagcgccttacttgaagttctcagttgtgacagaacactgltgattatag	2634
Db	2190	GAGCATAGACGGCTTACTGAAGTTTCTAGTTGATTCAGAAAGGGGAGGAGGACGACG	2249
QY	2635	aggaagcttatgaagatattca-----gcatactctcgtgaataaaca	2679
Db	2250	ACTATCTGGACCTGGAGAAATATTTCAGTGAAGACGACGACATTCACATCTGTCACA	2309
QY	2680	atgcacattgaacaagaagcttctccagaattccaagaacccctgagcctgagaaagc	2739
Db	2310	GTCGATTTGAACCAAGAAGCTTCTCCGAATTTCAAGACACCCTTAGCTAGGCAAAAGC	2369
QY	2740	aatttaatgcac-----2752	
Db	2370	AATTTAATGCAACCAATTTCCAGAAATAGATAGAGAAGACTGACCCTTGTTGCAC	2429
QY	2753	-----2752	
Db	2430	ACAGACACCTATGCTTAAATATCAAAATGTCTCTCTAGTGAATTTGTTGATGCTTGGC	2489
QY	2753	-----2752	
Db	2490	GACAGAGTCTACTCCACATGGGCTATCTTATCTGATCTCCAGAAAGCCAAATATAGA	2549
QY	2753	-----2752	
Db	2550	CTTTTCTGATGATCCATCACCCTGGAGCAATAGACAGTAAATACAGCCTGTCTGAAATGA	2609
QY	2753	-----2752	
Db	2610	CACACTTCAGGCCACAAGCTTCATCAACAAGTGGGAGACATGTATTTAACCCCTGAGTCAGCC	2669
QY	2753	-----2752	
Db	2670	TCCAATTAGATTAAATGAGAAACTGGGGACAATGCAGATCTCTGTGGGATTAAC	2729

[illegible]



D	3810	CTGGCTTAGTAATGGCTCAGAGATCAAGGAAFTGCATGGTATCTGCTCAGATCGGGACACA	3865
Q	3685	atgaaacalccatctcatcattccagtggaacatggttcaactgttacgaaaaaagg	3744
D	3870	ATGAAACATCCATTCTATTTCATTTCAGTGAGCATGTGTTCACTGATCAAGAAAAAGAGG	3929
Q	3745	agataaaagagacgtacaatctcatccaggtgttttggaaacagtggaatgttac	3804
D	3930	AGTATAAATAGGACCTGTACAATCTCTATCCAGGTGTTTGGAAACAGTGAATATGTAC	3989
Q	3805	catcaaaagtvgaaatttggcgggtggaaagccattatggcgagcatctacatctggga	3864
D	3990	CATCCAAAGCTGGATTTGGCGGGTGGAAATCCCTTATTGGGAGACATCTACATGCTGGGA	4049
Q	3865	tggacacacttttctggtgttacagcaataagtgatcagaactccccctggaaatggctctg	3924
D	4050	TGAGCACACTTTTCTGCTGTACAGCAATTAAGTGTACAGACTCCCTGGAAATGGCTCTG	4109
Q	3925	gaaacattagaagatlttcagatctacagcttcaggaaacaatatgaaacagtggccccaaagc	3984
D	4110	GAACACTTACAGATTTTTCAGATTACACTTTCAGGACATATAGGACAGTGGCCCCCAAGC	4169
Q	3985	tggccagactcaattatccggatccaatcaatggcttggaagacaaagagcccttctct	4044
D	4170	TGGCCAGACTTCAATTATTCGGATCAATTAAGCTTGAGACCAAGAGGCCCTTTTCTT	4229
Q	4045	ggatcaaggtggaatctggttgaccacaatgatalctacagcatcaagaaccagagtgccc	4104
D	4230	GGATCAAGGTGGATTCGTGGACACCAATGATTAATTAACGGCATCAAGCCAGGGTGCC	4289
Q	4105	gtcagaagttcccaagccttacaatctcctaagttatcaatcatgtatagctgtatggga	4164
D	4290	GTCAGAAGTCTCAGCGCTCTACATCTCTAGTTATCATCATGATTAATAGTTTATGGGA	4349
Q	4165	agaagtggcagacttatccggagaaattccactgtgaacctaaatgtgtctcttggcaatg	4224
D	4350	AGAGTGGCAGACTTATCGAGAAATTCACCTGGAACCTTAATGTGCTTCTTGGCAATG	4409
Q	4225	tggattcatctgggaataaacacaatatlttlaacctccaattatgtctcgatacaccc	4284
D	4410	TGGATTCACTGGGATTAACACAAATTTTTTAACCTCCAAATTATGTCTGATACATACC	4469
Q	4285	gtttgacccaactcatlataagcatctcgscagcactcttcgcgaatgaaatgtatgggtctgtg	4344
D	4470	GTTTGCACCCCACTCATTTATAGCATTCGCGACACTCTTCGGATGAGTGTATGGGCTGTG	4529
Q	4345	atttaaatgttgcagcatccacttggaaatgggaatlaagaaatcagatctcagatgcaga	4404
D	4530	ATTTAAATGTTGCAGATCCCATTTGGGAATGGAGAGTAAAGCAATACATCATCACAGA	4599
Q	4405	ttactctcatcctcaacttaccaatatggtttgcacctgtctccctcaaaagctcgac	4464
D	4590	TTACTCTCTTCATCTCACTTATACCAATATGTTTGGCACCCTGCTCTCTTCAAAAGCTGCAC	4649
Q	4465	ttcaactccaaggaggagtaatgctcgtggaacctcaagtgaaataatccaaaagatgyc	4524
D	4650	TTCACTCTCAAGAGGAGGAGTAATGCTGGAGACTCTCAGTGTAATTAATCCAAAAGATGGCC	4709
Q	4525	tgcgaatlggaacttccgaaagaacaaatgaaatgcagaagatgaacttaccatccagggatataat	4584
D	4710	TGCAATGTGGCTTCAGAAACAAATGAAGTCAACAGGAGTAATCACTCAGGAGTAATAAT	4769
Q	4585	ctctgtcttaccagcatgatgttaaagaagtcccatctccaggaatcaagaatggcatc	4644
D	4770	CTCTGCTTACCAACATGATATGTGAAGAGTTCCTCATCTCCAGCATCAAGATGGCCATC	4829
Q	4645	agtgaaactcctcttctcagaatgycaaagttaaagtttccagggaaatcaaaagctctc	4704
D	4830	AGTGAAGCTCTTTTTCAGAAATGGCAAGTAAGTTTTTCAGGGAATCAAGACTCTCT	4889
Q	4705	tcaacacctgtgtgaaactctcagaccacogttactgtaactcgttaacttgaatcaacc	4764

Db	4890	TCACACCTGTGGTGAACCTCTTAACACCACCCAGCATTACGACTGCCTACCTTGCAATTCACG	494
Qy	4765	cccaagatctgggtgcaccagaattgcccctgaagabtgaaagtctcgtggctcgagagacagg	4824
Db	4950	CCCAGAGTTGGGATGCACCAAGATTGCCCTGAGGAGATGAGGTTCTGGGCTCGAGGCACAGG	5009
Qy	4825	acctctaactgcagcagtc	4848
Db	5010	ACCTTACTGAGGGTGCGCACACTGC	5033
RESULT	9		
AR029098			
LOCUS	AR029098	4334 bp	DNA
DEFINITION	Sequence 38 from patent US 5859204.	PAT	29-SEP-1999
ACCESSION	AR029098		
VERSION	AR029098.1	GI:5941071	
KEYWORDS	.		
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	Unclassified.		
AUTHORS	1 (bases 1 to 4334)		
TITLE	Lollar,J.S.		
JOURNAL	Modified factor VIII		
FEATURES	Patent: US 5859204-A 38 12-JAN-1999;		
	Location/Qualifiers		
Source	1..4334	/organism="unknown"	
BASE COUNT	1109 a	1126 c	1085 g
ORIGIN			1014 t

[illegible]



Db	543	CACGTGACCTGGTGAAGAGCTGAAATTCGGGCTCATTTGGACCCCTGCTGGTTGTATA	602
Qy	1017	gaaggagctctgcgcgaagaaagacacagacctctgcacaattatactacttctgc	1076
Db	603	GAAGGAGATCTGACACAGAAAGAACCCCAACACTGCAGCAATTTCTACTTTTGT	662
Qy	1077	gtatttgatgaaggaaaglttgccactcsgaaacaagaacactccttgaatgagatag	1136
Db	663	GCTTTTGATGAGAGGAAAGTAATTTGGCACTCGCAAGAAGAAATGACTCTTGCGACAGGGCCATG	722
Qy	1137	gatctgcatctgtctcgggcttcggccttaaatgtcacacngtloahbgttatgtaacng	1196
Db	723	GATCCCGCACTGCGAGGGGCCACACCTCGCATCACACAGTCAATGCTATGTCAACAG	782
Qy	1197	tctctgcgaagtcgcatctgcatgcacagagaatacagtcatactatgcatgtgaatg	1256
Db	783	TCTCTGCCAGGTCTGATCGGATGTGCATAGAAATCATAGTCTACTGCGCAGCTGATTTGGAATG	842
Qy	1257	ggcaccaactctctgaagtgcaatcaataltcctcgaaggtacaacatltctltgtgaagac	1316
Db	843	GGCACACAGCCCGGAAGTGCACTCCATTTTCTTGAAGGCGACACGTTTCTCGTAGGCAC	902
Qy	1317	catcgccaggcgctcttggaatacctcgccaataacttctcctactcgtctaaacactctg	1376
Db	903	CATGCGCACAGGCTTCCCTTGGGAGATCTGCGCCACTAACCTTCTCACTGCTCAGCAATTCCTG	962
Qy	1377	atggaccttggaacagttctactcgtttttgcatatctctccacaacaaatgataagcgt	1436
Db	963	ATGACACTTTGGCCAGTTCCTACTACTGTTTGTTCATATCTTCCACACCACCATGCTGGCATG	1022
Qy	1437	gaagcttatgtccaagatgagaacgctgtgccagagaaaccccaactcagaatganaaatat	1496
Db	1023	GAGGCTACACGTCAGAGTAGAAGACGTGGCGCGAGAGACCCACAGCTGGGGAGAAACTGAT	1082
Qy	1497	gaagaagcgggaagactatgtatgtactactgattctgaaatggaatgagtgtcagtltt	1556
Db	1083	GAAAGA---GGAAAGATTTGATGACAATTTGTACAGATCGACATGGACGTGGTCCGCGCTC	1139
Qy	1557	gatbatbaaacactctctctctatccaaatctcgctagttgcgaagagaatctctaa	1616
Db	1140	GATGCTGACGACGTTCTCCCTTTATTCCAATCCGCTCGGTTCGCAAGAAAGATCCCAA	1199
Qy	1617	acttgggtacattcaatctgtctctgtagaagagagagacttggagactatgctctctagctc	1678
Db	1200	ACCTGGGGCACATCAATCCTCTGCAGAGAGAGAGACTGGGACTACGGCCCCGGGCTCCC	1255
Qy	1677	ggccccgcatgacagaagtatlaaagtcaatatltgaaacaatgycctcctcagcagatg	1738
Db	1260	AGCCCCAGTGACAGAAAGTTTAAAGCTCTACTTAAACAGTGGCTTCACGCAATTTGGT	1319
Qy	1737	aggaagatcaaaaaaagtcggatcttgaatgaacagatggaacacttlaagactctgtgaa	1796
Db	1320	AGGAATAATCAAAAAAGCTCATTCGTCTGCTACCGAGATGTAACTTTAAAGCTCTGTAA	1799
Qy	1797	gcatctacagcatgaatcaagaaatcttgggaaccttcttacttggggaagtgtgaacaca	1856
Db	1380	GCTATTCGATATGATATAGGAATCCCTGGGACCTTTACTTATGAGAAAGTTGAGACACA	1439
Qy	1857	ctgttgattatattaaagaaacaagcagaagcaatataacatctaacctcctcagyaatc	1916
Db	1440	CTTTTGGATTATTTAAGAAATTAAGGAGGCGCCACTATATACATCTACCCATCATGATTC	1499
Qy	1917	actgaibtcgctcttctgttatcttaagagatataccaaaagtgataaacatttgaagat	1976
Db	1500	ACTATATGTCACGGCTTTCGACCCAGGAGACTTCTTAAAAAGTTTGAAGCAATTTGAAAGAC	1559
Qy	1977	tctcaactctgcgcagagaaatcatcaataataatgagacgtactgtgaagaatg	2036
Db	1560	ATGCGCAATTCGCCAGAGAGAGACTTTCAAGTATTAATATGACACGTGACTGTGGAAGATGG	1619
Qy	2037	ccaactaatcagatcctcggtygcctgaccgcgtactactagtttcgttaataatgag	2096

Db	1620	CCACCAAGTCCGATTCCTCGGTGCTTGACCCCGCTACTACTCGAGCTCCATTAACTAGAG	1679
Qy	2097	agagatcagcttcaaggacatcatgagccctccatctcatctgttacaagaatctgtatag	2156
Db	1680	AAAGATCTGGCTTGGGACATCATTTGGCCCTCTCTCATCTGCTACAGAAGATCTGTAAAC	1739
Qy	2157	caaaagagaaacacagataatgttcaagacaagaagaaatgtcatccgttcttctgtatattg	2216
Db	1740	CAAAAGAGGAAACACAGATGATGTACAGCAAGAAAGAAAGCATCTCTTTTCTGTATTGCAT	1799
Qy	2217	gagaaacgaagctgtgtacccccaagaagaataatacaacgcttctcccaatccagctgtga	2276
Db	1800	GAGATTCAAACCTGTGTCTCTCGCAGAGAAATATTTCAACGGCTTCTCCCAATCCGATGGA	1859
Qy	2277	gtgcagctttaggagatcccaagagttccaagacctccaacatcatgacaagcaatgaatgtcat	2336
Db	1860	TTACAGCCCCAGGATCCAGAGTTTCCAAAGCTTCTTAACATATGACACAGCATTCATGTGCTAT	1919
Qy	2337	gtttttgatatgtttgtcagctgttgcagtttgtttgtcaltgagtgagcaactgtgtatctcta	2366
Db	1920	GTTTWTGTACTTCTGACAGCTGTGCGTTGTGTTTGCACAGAGGTGACATCTGTATCTTA	1979
Qy	2397	agcaattgagacagacagactgactctccctctctgtctctctctctctgtgataactcaaac	2456
Db	1980	AGTGTGGAGACAGACGGAGCTTCTCTCCGTCTTCTCTCGGTATCACCTCTTCAACAC	2039
Qy	2457	aaaaatgtctcta tgaagacacacacacccaattcccaattccaagagaaacttctctga	2516
Db	2040	AAAAATGCTCTATGAGACACACTACCTCTGTTCCCTTCTCAGGAGAAACGGTCTTTCATG	2099
Qy	2517	tcgaatgaaaaaaccccaagttctctatgtatctgtgggtgcccacaacttcagacttctgagaacga	2576
Db	2100	TCAATGGAACAAACCCAGGCTCTCTGGGTCTATGGGTGCGCACACATCGACTTCGGGAAACGA	2159
Qy	2577	ggcaatgacccgcttactcgaagagtttctagtgtgtgaacaagaacactgtgtatattacag	2636
Db	2160	GGGATGACAGGCTTACGTGAAGAGGTATATGTTGTGACAGGAGACATTGTGTATTTATATAC	2219
Qy	2637	gacagttatgaagaatatttcaagataacttgctgtgagtaaaacaatgtgccaattgaaccaaga	2696
Db	2220	AACACTTATGAAGTATTTCCAGGCTTCTTGCTGTAGTGGAAAGATGTCAATTGAACCCAGA	2279
Qy	2697	agctcttcccaagaatttcaagaacacctlagactagagcaaaagcaattatattgcacccca	2756
Db	2280	-----	2279
Qy	2757	ccagatcttgaaacgcacataacgsgaataaactcgtactactcttcagtcaagatacaag	2816
Db	2280	-----GACATTAAAGGCTTCTCTTTCACGCGGAGGAAAC	2315
Qy	2817	gaaatbactatbtatgttaccataatgttgaatgaagaagaagaatttgaacattat	2876
Db	2316	AAAAATGACTATGTATGTATATCTTCTCACTACGAAACCAAGGAGAAATTTTGACATTTTAC	2375
Qy	2877	gatatagatgaaaatctcaagacccccgcgcgaacttccaagaagaacacgcgaactatttlat	2936
Db	2376	GGTAGAGATGAATAATCGAGACCTCCGACACTTTCACAAGACAGACCGACACTATTTCATT	2435
Qy	2937	gtctgcagtgagaagagctctgtggaattatgtgatgtagttaagctccccaatgttctctaagaac	2996
Db	2436	GCTTCGGTGTGAGCGACTCTGTGGATATACGGGATGAGCGAATCCCCCGGGCGCTTAAAGAAC	2495
Qy	2997	aggagctaaagtgtagtgcttcacctgaagttcaagaagaatgttttccagaaatttaccgat	3056
Db	2496	AGGGCTCAGAACGGAGAGGTGCTCTCGGTTCAAGAAAGTGTCTTCCGGGAATTTGTCTAC	2555
Qy	3057	ggccctcttaacaccccttataacgttgagaaacaaatgaacatttggagactccgtgg	3116
Db	2556	GGCTCTTACACGACGCGCTGTATCCCGGGGAACTCAACAAACACTTGGGGCTCTTGGA	2615
Qy	3117	ccatatataaagacgaagagtttgaagataatcatcatgttgaacttccaagaatcgaagcctct	3176
Db	2616	CCCTACATCAGAGCGGAAGTTGAAGACAACATCATGTGTAACTTTCAAAACCCAGGCGCT	2675



QY	3177	cgccccattccctctctactctagacctatctctatctatgaggaagatccgaaggccaagagca	3236
Db	2676	CGTCCCTATTCTCTTACTGTGAGGCTTAATTTCTTATCCGGATGATCAGACAAAGGGCCA	27353
QY	3237	gaaccatgaaaaaaactctgtccaagccaaatgaaaccaaaacttactcttctgaaagtcca	3296
Db	2736	GAACCTCGAACAAACTCTGCTCCACCAAAATGAAACCAAGAACTTCTTGTGAAAGTGACG	27959
QY	3297	catcatatgagccaccactaaagatgagtttgactgcgaagccctgggcttacttctctgat	3366
Db	2796	CATCAATGTGACACCACAGAAAGAGATTTTGACTGCAAAAGCCTGGCCCTACTTTTTCAT	28555
QY	3357	gttgacctggaaaaaagaatgtcactaaaggcttgatctggaacccctctgtgtctgcacaact	3416
Db	2856	GTTGACCTGGAAAAAGATGTGCACTCAGGCTTGATCGGCCCTTCTGTATGTGCGGGCC	2915
QY	3417	aacacacatgaaacctgtctcatgaggagaacagttagacagtacagaagaattgtctcttttc	3470
Db	2916	AACACCTGAAACCGCTGCTCAGGGTAGACAGATGACCGTGCAGAAATTTGCTCTGTTTTC	2975
QY	3477	accactcttgaatgagaccaaaagctgttacttcaactcgtaaaaatctatgaaagaactgcag	356
Db	2976	ACTATTTTTGATATAGACAAAGAGCTGTACTTCTCTGATAAATGTGGAAAGGAACTGCCGG	3035
QY	3537	gtccccctgcaatatcccaatgtaagaatcccaactcttaagaagaatttgcttccatgca	3566
Db	3036	GCCCCCTGCCACTGCAGATGAGAGGCCCACTGTAAAGAAACTATGCTTCCATATCA	3095
QY	3597	atcaatgctctacataatgataacactcctgtcttagtaatgtgtctcaagatccaagaagt	3656
Db	3096	ATCAATGGCTATGTGATGGATACCTCCCTGTGCTAGTAATGGCTGCAGAAATCAAGATC	3155
QY	3657	cgaatgatactgctcgaatctggcagcaatgaaacatccaattctatctatctcaatgta	3716
Db	3156	CGATGGTATCTGCTCAGCAATGGGACGAATGAATATTCATTCGATTCATTTTATGGCGA	3215
QY	3717	catgtttcaactgtacgaaaaaaagaagagatataaaatgcaactgtcaactctatcca	3776
Db	3216	CACGTCTTCAGTGTACCGAAAAAGAGAGATATATAATGGCCGTGTCAATCTTATCTCCG	3275
QY	3777	ggtgtctcttgagaacagtgaagaatgttcccatcccaagcttgaaattctggcgggtgaaatgc	3836
Db	3276	GGGTCTTTGAGACAGTGAATGTCTACCGTCCGCAAAAGTTGGAATTTGGCAATGAATGC	3335
QY	3837	cttatgtggagagatactcaactgtctgggaatgaagcaacttcttcgtgttaagcaataag	386
Db	3336	CTGATTGTGGGAGACCTGTCAAGCTGGGATAGACACGACTTTCCGTGCTACGCAAGAG	3395
QY	3897	tgtcgaactcccccctgggaatgtgtctctcggacacatagagatttccagatctaacgcttca	3956
Db	3396	TGTCAAGCTTCCACTGGGAATGGCTTCTGGACGCTTAGAGATTTTTCAGATCACAGCTTCA	3455
QY	3957	ggacaaatagagacagtggccccaaagcttggccagacttcatatctccgaatccaat	4016
Db	3456	GGACATATGAGACAGTGGGCCCAAAAGCTGGCAGACTTATTTATTCGGATCAATCAAT	3515
QY	4017	gctctggaagcacaagaagcccttctcttgatacaaggatgatctgttggcaaccaatgat	4076
Db	3516	GGCTGGAGACCAAGAGATCCCACTCCTGTGATCAAGGTGATCTGTTGGCACCATATGTC	3575
QY	4077	attcagaagcaatcaagaagccaggatgcccgtcagaagtctccagccctctcatctccag	4136
Db	3576	ATTCAAGGCATCATGACCAAGGATGCCGTCAGAGATTTTCCAGCCTTCAATCTCCAG	3635
QY	4137	tttatcatcatgatatgactcttgaatgggaagaagtgtagacatctacgaagaatttccact	4196
Db	3636	TTTATCATCATGATCAAGTCTTGACGGGAGGACATCGCAGAGATTACGAGGAATTCACG	3695
QY	4197	ggaaccttaatggtctctcttggcaatgtagatcatctgggataaaacaacaatattttt	4266
Db	3696	GGCACCTTATATGCTCTTCTTTGGCAATGTGACGCAATCTGGGATTTAAACCAATATTTT	3755

QY	4257	aaacctcaattatgtgcgaataaccggttgaccacaactatataagttcgagc	4316
Db	3756	AACCTTCGATTGTGGCTCGGTACATCCGTTTCACCCACATTAACGATCCGCGAGC	3815
QY	4317	actcttcgcattgagatttgatbgtgcgtgatttaataagtltagcatlbgcattggaatg	4376
Db	3816	ACTCTGCCAATGGAGTTGATGGCGCTGTGATTTAACAGTTGGAGCATGCCCTGGGAATG	3875
QY	4377	gagagtaagaacataatcagabgacagattactgtcttcacttccacttacaatatgttt	4436
Db	3876	CAGAAATAAGCGAATATGAGACTCACAGATCACGGCTCCCTCCACCTAAGCAATATATTT	3935
QY	4437	ggcaactgtgtctctttaaagaagtcggaattcacttaccctccaaggagagatgaahtgcgga	4496
Db	3936	GCACACTGTGTTCTCTTCAACAAGCCCGCATTCACCTCCAGGGCGGACGAATGCTCGGGA	3995
QY	4497	cctcagatgaaataaccaaaagatgtgtcgaatgagttccagaagaacatgaagaatc	4556
Db	3996	CCCCGGGTGACACACGCGAGAGAAATGGCTCGAGTGGACTCGAGAAAGCGTGAAGTTC	4055
QY	4557	acagagtaacttactccaaggagataaatactctctgcttaccagaatgatalgtgaagatlc	4616
Db	4056	ACAGGCAATCACACCCAGGGCGTGAAGATCCCTGCTCAGACAGATGATGTGAAGAGTTTC	4115
QY	4617	ctcactctccagcagtcagaatgacctcaatgtgactctcttttttcagaatgycagaagta	4676
Db	4116	CTCGTGTCCATTAATCTAGCGAGCGCGCGCTGGACCCCTTTTCTTCAGAGCGCACACG	4175
QY	4677	aaggtttttcagggaataatgaagctctcttcacacccgtgtgaaactctctgacccacgg	4736
Db	4176	AAGGTTTTCTAGGGCAATTCAGAGATCTCTCACCCCGGTGTGAAGCTGTGGACCCCGC	4235
QY	4737	ttactgactctctacctctgaatttcaaccccaagatcttggttgaccagatltgcctgag	4796
Db	4236	CTGTTCACAGCGCTACGTAAGAGATGCACCCACAGAGCTGGGGCGACACATGGCCTGAGG	4295
QY	4797	atggaagttcttggtgcgcaggaacagagactctactga 4835	
Db	4296	CTCGAGGTTCTAGAGATGTGAGCGCACAGAGATCTTCTAATGA 4334	

	RESULT	10		
CY089927	LOCUS			
CYU89927	DEFINITION	6786 bp	DNA	circular
	ACCESSION	U89927		SYN
	VERSION	U89927.1		
	KEYWORDS	GI:2190694		
	SOURCE	.		
	ORGANISM	unidentified cloning vector.		
	REFERENCE	artificial sequence; vectors.		
AUTHORS	TITLE	1 (bases 1 to 6786).		
JOURNAL	REFERENCE	Holtz,A. and Lou,Y.		
	AUTHORS	Unpublished		
	TITLE	2 (bases 1 to 6786)		
	JOURNAL	Holtz,A. and Lou,Y.		
COMMENT		Direct Submission		
		Submitted (18-FEB-1997) CLONTECH Laboratories, Inc., 1020 East Meadow Circle, Palo Alto, CA 94303-4230, USA		
		This vector can be obtained from CLONTECH Laboratories, Inc., 1020 East Meadow Circle, Palo Alto, CA 94303-4230, USA. To place an order, call (415) 424-8222 or (800) 662-2566, extension 1.		
		International customers, please contact your local distributor or subsidiary. For technical information, call (415) 424-8222 or (800) 662-2566, extension 3.		
		This sequence has been compiled from information in the sequence databases, published literature and other sources, together with partial sequences obtained by CLONTECH. If you suspect there is an error in this sequence, please contact CLONTECH's Technical Support Department at (415) 424-8222 or (800) 662-2566, extension 3 or E-mail TECH@CLONTECH.COM.		
		Location/Qualifiers		



```
source      1. .6786
            /organism="unidentified cloning vector"
            /db_xref="taxon:45196"
gene        152. .811
            /gene="HIS3"
            152. .811
CDS         /gene="HIS3"
            /codon_start=1
            /product="HIS3"
            /protein_id="AAB64382.1"
            /db_xref="GI:2190693"
            /translation="MTEOKALVNRITNETKIQIAISLKGSPLIAIEHSIPEKEAEVA
EQATQSOVINVHTIGFLDMLHALAKHSGMSLIVICGLHIDHIDHETEDCGIALGO
AFKELALRGVRRFGSGFAPALDEALSRAYVDLSNRPYAVVELSIOREKVDLSCEMP
HLESEFAEASRIPLHYDCLRGKNDHHRSESAFALAVAIREATSPGNTDVPSTKVL
M"
gene        2729. .3532
            /gene="URA3"
            2729. .3532
CDS         /gene="URA3"
            /codon_start=1
            /product="URA3"
            /protein_id="AAB64383.1"
            /db_xref="GI:2190696"
            /translation="MSKATYKERATPSPFAKLFINHEKOTNLCASIDVRYTEL
LEVEALGPRIKLTHTVDILTFDESGEYVYKPLKALSAYNLFLEPDRPADIGNTVK
LOYSAGVYRIAEADITNAHGVYPCIVSGIKOAEVEMKEPGLIMLAELSGKSL
TGEYTKGYVDIAKSDKDFVIQFIQRDMGSRDGYMLINTPEVGLDDKDALGOQYR
TYDDVYSTDSDIIVTGRGLFAKGRDAKVEGERYRKAGMEAYLKRCSQON"
gene        4846. .5706
            /gene="AmpR"
            complement(4846. .5706)
            /function="ampicillin resistance"
            /codon_start=1
            /product="beta-lactamase"
            /protein_id="AAB64384.1"
            /db_xref="GI:2190697"
            /translation="MSIOHFRVALIPFPAFLCPVPAHPETLVKVADEDOLGARVY
IELDNSGKILIESFRPEFRPMSTFKVLGAVISRIDAGOBOLGRRIHYXNDNVE
YSPYTKRHLDGMYRELCSAATYMSDNTRANILITIGPKRLTFLNMGHVRVL
DRNEPLNEATIPDERDITPAVMATTLKRLGELTLASRQOLLDMWEADVAPVL
LRSAIPAGWFIADKSGAGERSKGIITAAALGPDKPSRIYIYTTGSOATMDERNQIA
EIGASLIIKHM"
BASE COUNT 1782 a 1590 c 1749 g 1665 t
ORIGIN
Query Match      35.6%; Score 2830.6; DB 56; Length 6786;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2836; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 5091 ccagctcatatgaatcgcccaacgcgcgggagagagcggttcgcatattggcgctc 5150
      |||||||
DB 3847 CCAGCTCATATTAAATGAAATCGGCCCAACGCGGGAGAGCGGTTCGTATGGGCCCTC 3906
QY 5151 ttccgctctcgcctcactcactcgcgtcgcgtcgcgttcgcgttcgcggagagcgatc 5210
      |||||||
DB 3907 TTCGCTTCTCCTCGCTCACTACTCGCTGCGCTCGCTTCGCTGCGCTGCGGAGCGGATC 3966
QY 5211 agctcaactcaaaagcggttaactcgtatccacagaaatcaggggataacgcaggaagaa 5270
      |||||||
DB 3967 AGCTCACTCAAAAGCGGATTAACGGTTATCCACAGATCAAGGAGATTAACGACGAAGAA 4026
QY 5271 catgtgagcaaaagcgcaagaaagcgcaagaaacgtaaaaagcgcgcttgcgtgcgtc 5330
      |||||||
DB 4027 CATGTGAGCAAAAGCGCAGCAAGGACCGTAAGGAGCGCCGCTGTGCGGCTT 4086
QY 5331 ttccatagagctccgcccctcagagcattcacaaaaatcgagagctccaagtcagagtg 5390
      |||||||
DB 4087 TTTCATAGAGCTCCGCCCTGACGAGCATCACAAAAATCGAGCTCAAGTCAAGAGTG 4146
QY 5391 gcgaaacgcagagactaaagataccagcgcttcccccctggaagctccctcgtgcg 5450
      |||||||
```

```
DB 4147 GCGAAACCCGACAGACTATTAAGATACAGCGCTTCCCTGGAAAGTCCTCGTGGC 4206
QY 5451 ctctccctgttcgagccctgcgcgttacccgataaccttcgcgcttctcccttcgggaag 5510
      |||||||
DB 4207 CTCCTCGTTCCGACCCCTGCGCTTACCGGATACCTGCTCCGCTTCCCTTCGGGAG 4266
QY 5511 cgtgcgcttctcaatgctcagcgtcgttaggtatctcgaatcgtcgtgtaggtcgtc 5570
      |||||||
DB 4267 CGTGGCCCTTTCTATAGCTACGCTAGGTATCTCAGTTGGGTAGGCTTCGCTC 4326
QY 5571 caagctggcgtgtgtgcaggaacccccgcgttcagcccgacgcgctgcgtctatccgtaa 5630
      |||||||
DB 4327 CAACCTGGCGCTGTGACAGCAACCCCGCTTCAGCCGACCGCGCTGCGCTTATCCGTTAA 4386
QY 5631 ctatcgctctgagtcgaaccccggttaagaacagacttatcgccatctgagcagacgaacttg 5690
      |||||||
DB 4387 CTATCGTCTTGAGTCCAAACCCGCTTAAGACACAGACTTATCCGACTGCGACAGCCACTGG 4446
QY 5691 taacagattagcagagcagagtatgtagcggtgtcacaagctcttgaaagtgtgagc 5750
      |||||||
DB 4447 TAACAGGATTAGCAGACGAGAGCTATGTAGGGGTGCTACAGAGTTCTTGAAGTGTGGCC 4506
QY 5751 taactagcgtcacactagaagagcagatttgytatctgcgtctgtgaaagccaattac 5810
      |||||||
DB 4507 TAACTAGGCTACACTAGAAGACAGATTTGTATCTGCGCTCTGCTGAAGCCAGTTAC 4566
QY 5811 cttcggaaaaagatgtgtgctctgtatccgcggaacaaacccgcgtgtagcggtg 5870
      |||||||
DB 4567 CTTCGGAAGAAAGAGTGTGATCTTATCGGCAACCAACCCGCTGTACCGGTGG 4626
QY 5871 ttttttgcttgcagagcagcagattacgcgcgaagaaaaagagatctcaagaagatccctt 5930
      |||||||
DB 4627 TTTTGTGTTTGCAAGCAGCAGATTAACGCGCAGAAAAAGATCTCAAGAAATCCTTT 4686
QY 5931 gatcttctcagcgggtctcagcgtcagctcagctggaagaaactcacgcttaaggaatttgct 5990
      |||||||
DB 4687 GATCTTTTCAAGGGGTCTGACGCTCAGTGAAGAAACACAGTTAAGGAGTTTGGT 4746
QY 5991 catgagattatcaaaaagatcttcacctatgatccctttaaataaaatgaaatttaa 6050
      |||||||
DB 4747 CATGAGATTATCAAAAAGAGATCTTACCTTGATCTTAAATTAATAAAGAACTTTTAA 4806
QY 6051 atcaatctaaagtatatatagtaaaacttgctgcagcagttaccaaigtcaactcagtgta 6110
      |||||||
DB 4807 ATCATCTAAAGTATATATAGTAAGTAACTTGGTCTGACAGATTACCAATGCTTAAATCA 4866
QY 6111 ggcacctatctcagcagatctcgtctattcgtlcatccatagttgcctgactcccgctcgt 6170
      |||||||
DB 4867 GGCACCTATCTCAGCGCATCTGTCTATTCTGTCATTCATAGTTGCTGACATCCCGTCTG 4926
QY 6171 gtgataactcagatacaggaagggtcttaccatctggcccagtgctgcaatgataccgcg 6230
      |||||||
DB 4927 GTAGATTAACATACGATACGGGAGGGCTTACATCTGAGCCCAAGTCTCAATGATACCGCG 4986
QY 6231 agagccacgcctcaccgcgctcagaattatcagcaataaacacagcagcggaaaggcgca 6290
      |||||||
DB 4987 AGACCCACGCTCACCGCTCCAGATTATACGCAATTAACACGACGACCGGAAGGCGCA 5046
QY 6291 ggcgagaagtgtctcgcgaacttatccgctccatccatccatctatattgttgcgga 6350
      |||||||
DB 5047 GCGCAGAAAGTGTCGCTCAACTTATCCGCTCCATCCAGTCATTAATGTGTCGGGA 5106
QY 6351 agcttagagtagtagtctgcgcagttatatttgcgcaagctgtgtgccaattgctaaagg 6410
      |||||||
DB 5107 AGCTTAGAGTAGTAGTTCGCGCAGTAATTAATTTGCGGAAGCGTGTTCATTCGTTAAGG 5166
QY 6411 catcggtgtcagcgtcgtcgttggtaaggcttaataagctccoggttcccaagatc 6470
      |||||||
DB 5167 CATCGTGTGTCACGCTCGTCTGTTGATGCTTCACTTACGCTCCGGTTCACACATC 5226
QY 6471 aagcgagttacatgataccccaatgtgtgtgcaaaaaagcggttagctcctcgtctcc 6530
      |||||||
DB 5227 AAGCGAGTTACATGATGCCCATGTGTGTGCAAAAAGCGGTTAAGTCTTCGGTCTCC 5286
```



QY 6531 gatcgttgcagaaagtaagtttggccgagtgatcaactcaatggtatgcaagcactgca 6590  
 |||||||  
 Db 5287 gatcgttgcagaaagtaagtttggccgagtgatcaactcaatggtatgcaagcactgca 5346  
 |||||||  
 QY 6591 taattctctactgtcatgcaatccgtaagatgcttctgtgactggtgtaactaac 6650  
 |||||||  
 Db 5347 taaattcttctactgtcatgcaatccgtaagatgcttctgtgactggtgtaactaac 5406  
 |||||||  
 QY 6651 caagtcaactctgagaatgatagtgagcgagcttgccttgcgagcggtcaataag 6710  
 |||||||  
 Db 5407 caatgcatcttgcagantagtgatgagcgagcttgccttgcgagcggtcaataag 5466  
 |||||||  
 QY 6711 ggaataatccgagcagacaagaagaacttaaaagtgcatacatggaanaagcttctc 6770  
 |||||||  
 Db 5467 ggataatattccgagcagacacatagagaacttaaaagtgcatacatggaanaagcttctc 5526  
 |||||||  
 QY 6771 gggagcgaataactctcaagagatcttgcgctgtgtgagatccagttcgatgtaaccactcg 6830  
 |||||||  
 Db 5527 gggcggaataactctcaagagatcttgcgctgtgtgagatccagttcgatgtaaccactcg 5586  
 |||||||  
 QY 6831 tgcacccaactgacatcttcaagatcttcaacttcaacagcgcttctgagtgagcaaaaac 6890  
 |||||||  
 Db 5587 tgcacccaactgacatcttcaagatcttcaacttcaacagcgcttctgagtgagcaaaaac 5646  
 |||||||  
 QY 6891 aggaaggaacaaatgcgcgcaaaaaggaataagagcgagcaacggaatgltgaatactcat 6950  
 |||||||  
 Db 5647 aggaaggaacaaatgcgcgcaaaaaggaataagagcgagcaacggaatgltgaatactcat 5706  
 |||||||  
 QY 6951 actcttcttctcaataatattatgaagcattatcaaggttatgtctcaatgagcgata 7010  
 |||||||  
 Db 5707 actcttcttctcaataatattatgaagcattatcaaggttatgtctcaatgagcgata 5766  
 |||||||  
 QY 7011 cataattgaaatgtatttgaagaaataacaaatagagggttccgcgcacatttcccgaa 7070  
 |||||||  
 Db 5767 catatttgaatgtratttttgaagaaataacaaatagagggttccgcgcacatttcccgaa 5826  
 |||||||  
 QY 7071 agtgcacactlgaagctlctaaagaacattatatacaatgaatcaacttaaaatagagc 7130  
 |||||||  
 Db 5827 agtgcacactlgaagctlctaaagaacattatatacaatgaatcaacttaaaatagagc 5886  
 |||||||  
 QY 7131 tatacgaagcgcttctgcctgcgagcttgcggtgataagcggtgaanaacacttgcacat 7190  
 |||||||  
 Db 5887 tatcaccagagcgcttctgcctgcgagcttgcggtgataagcggtgaanaacacttgcacat 5946  
 |||||||  
 QY 7191 gcaagctcccgagagagcggtcacaagcttgcgtgaagcgagatgcggaagcaagaacg 7250  
 |||||||  
 Db 5947 gcaagctcccgagagagcggtcacaagcttgcgtgaagcgagatgcggaagcaagaacg 6006  
 |||||||  
 QY 7251 tcaagcgcgctcagcgaggtgttgcgaggtgtcgcgagctgacttaactatgcgagca 7310  
 |||||||  
 Db 6007 tcaagcgcgctcagcgaggtgttgcgaggtgtcgcgagctgacttaactatgcgagca 6066  
 |||||||  
 QY 7311 gcaagcttgcctgagagagcgacatatagtcgtgtgaaatccgacaaatggttaagag 7370  
 |||||||  
 Db 6067 gcaagcttgcctgagagagcgacatatagtcgtgtgaaatccgacaaatggttaagag 6126  
 |||||||  
 QY 7371 aaaaatccgacatcagaagaatgttaaacgtttaaatcttgcgttaaatctt 7430  
 |||||||  
 Db 6127 aaaaatccgacatcagaagaatgttaaacgtttaaatcttgcgttaaatctt 6186  
 |||||||  
 QY 7431 tgttaaatcagctcaattttaaacaatagcggaatccggaanaatccctataatca 7490  
 |||||||  
 Db 6187 tgttaaatcagctcaattttaaacaatagcggaatccggaanaatccctataatca 6246  
 |||||||  
 QY 7491 aaagaatagacggagataaggttgatgtgttccagtttggaaagaagtcactatla 7550  
 |||||||  
 Db 6247 aaagaatagacggagataaggttgatgtgttccagtttggaaagaagtcactatla 6306  
 |||||||  
 QY 7551 aaagaatagacggagataaggttgatgtgttccagtttggaaagaagtcactatla 7610  
 |||||||  
 Db 6307 aaagaatagacggagataaggttgatgtgttccagtttggaaagaagtcactatla 6366  
 |||||||

QY 7611 cgtgaaccatcacccatcaatgatttctgggtcgaagtgccgtaaaccaataatcgg 7670  
 |||||||  
 Db 6367 cgtgaaccatcacccatcaatgatttctgggtcgaagtgccgtaaaccaataatcgg 6426  
 |||||||  
 QY 7671 aacctaaagggagcccccagatttagagcttgacgggggaaagccggcgaaacgttgagaga 7730  
 |||||||  
 Db 6427 aacctaaagggagcccccagatttagagcttgacgggggaaagccggcgaaacgttgagaga 6486  
 |||||||  
 QY 7731 aaggaaggaagaagcgaagaagcgagcgctgaagcgctggaagtgtagcgtaacg 7790  
 |||||||  
 Db 6487 aaggaaggaagaagcgaagaagcgagcgctgaagcgctggaagtgtagcgtaacg 6546  
 |||||||  
 QY 7791 ctgcgctgaataccacacaccccgcgcttaatgctgcgcgtacagggcgctgcgcat 7850  
 |||||||  
 Db 6547 ctgcgctgaataccacacaccccgcgcttaatgctgcgcgtacagggcgctgcgcat 6606  
 |||||||  
 QY 7851 tgcgcatcagctagcgaacatgcttggaagggagatggtgagcgctctctgatatla 7910  
 |||||||  
 Db 6607 tgcgcatcagctagcgaacatgcttggaagggagatggtgagcgctctctgatatla 6666  
 |||||||  
 QY 7911 cgcagctgctgcagggggggggg 7935  
 |||||||  
 Db 6667 cgcagctgctgcagggggggggg 6691  
 |||||||

RESULT 11  
 CVGEM13ZP  
 LOCUS  
 DEFINITION Cloning vector pGEM-13zf(+).  
 ACCESSION X65315.2  
 VERSION GI:5420149  
 KEYWORDS beta-lactamase; bla gene; cloning vector; lacZ gene; multiple cloning site; phase I region; promoter; pUC/M13 primer binding site.  
 SOURCE  
 ORGANISM  
 Cloning vector pGEM-13zf(+).  
 Cloning vector pGEM-13zf(+).  
 Official sequence; vectors.  
 1 (bases 1 to 3179)  
 Technical Services.  
 Direct Submission  
 Submitted (23-MAR-1992) Technical Services, Promega Corporation, 2800 Woods Hollow Road, Madison, WI 53711-5399, USA  
 JOURNAL  
 REMARK  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 Submitted (28-MAY-1993) Technical Services, Promega Corporation, 2800 Woods Hollow Road, Madison, WI 53711-5399, USA  
 revised by [3]  
 3 (bases 1 to 3179)  
 Kendrick, M.  
 Direct Submission  
 Submitted (02-JUL-1999) Kendrick M., Technical Writing, Promega Corporation, 2800 Woods Hollow Road, Madison, WI 53711, USA  
 On Jul 8, 1999 this sequence version replaced gi:58170.  
 See X65300-X65335 for related vector sequences  
 This vector can be obtained from Promega Corporation, Madison, WI.  
 Call one of the following numbers for order or technical information:  
 Order or Technical 800-356-9526  
 In Wisconsin 800-356-9526  
 Outside U.S. 608-274-4330.  
 FEATURES  
 source  
 1. 3179  
 location/Qualifiers  
 /organism="Cloning vector pGEM-13zf(+)"  
 /db\_xref="taxon:90121"  
 misc\_feature  
 1  
 /note="T7 transcription initiation site"  
 11..43  
 /note="multiple cloning sites"  
 49..68  
 /note="Sp6 promoter"  
 misc\_feature  
 51







QY 6651 caagtcattctgagaatgtagtgcgagcagagctgctctgcgcgcgtcaatag 6710  
 1808 CAAGTCATTCTGAGATATGATGCGGACGAGTGTGCTGTGCGCGCGTCATPAG 1867  
 QY 6711 gqataatccgcgcacacatagcagaactttaaagtctcattcaattggaacgctcttc 6770  
 1868 GGATATATCCCGCGCACATAGCAGAACTTTAAAGTCTCATCTATTGGAAGAAAGCTTCTTC 1927  
 QY 6771 ggggcaaaactctcaagaactctaccgctgtctgagatccagtctgaattaaaccactcg 6830  
 1928 GGGCGGAAACTCTCAAGGATCTTACCGCTGTGAGATCCAGTTGCATGTAAACCACTCG 1987  
 QY 6831 tgcaccaactgatactcagcactcttactctcacaacgctctccgggtgagcaaaac 6890  
 1988 TGCACCCAACTGATCTTGAGCATCTTTTACTTTCACACAGCTTTCTGGGTGAGCAAAAC 2047  
 QY 6891 aggaagcacaatgcgcgcaaaaggaataagcgacacaggaatgttgaaactcat 6950  
 2048 AGGAAGGCAAAATGCCGMAAAGGAATAGGCGCACACGAAATGTTGATACTCAT 2107  
 QY 6951 actctctcttctcaatatattatgaagcattatcagggcttctcattgagcgata 7010  
 2108 ACTCTCTCTTTTCAATATTATTGAAGCATTTATCGGGTATGTCTCATGAGCGGATA 2167  
 QY 7011 catattgaatgtaattagaanaataaacaataggggtctccgcgcacattcccgcaaa 7070  
 2168 CATATTGAATGATATTGAAAAATAAACAATAGGCGTCCGCGCACATTTCCCGGAAA 2227  
 QY 7071 agtgcacactgagctcctaagaacatactatcatgacataactataaataagcg 7130  
 2228 AGTGCACACTGACGCTCAAGAAACCATATATATGACATTAACTATAAAATTTGGGCG 2287  
 QY 7131 taccacgagccctctgctcgcgcgttgcgtgatagcggtgaaacactctgaacat 7190  
 2288 TATCAGAGGCGCTTCTGCTGCGCGTTCGCTGATGACGGTGAACACTCTGACACAT 2347  
 QY 7191 gcaagctccgcgagacgctcagctctgctgaagcgagatgcgcgagacagaacgccc 7250  
 2348 GCAGCTCCCGGAGACGCTGACAGCTTGTCTGAAGGAGATGCCGGAGACAGAACCCCG 2407  
 QY 7251 tcagagcgcgctcagcggtgtgtgcggtgtctgcggtctgcttaactaagcgacata 7310  
 2408 TCAGGCGCGCTCAGCGGCTGTTGGCGGCTGTGGGCGCTTAATATGCGGCATCAAA 2467  
 QY 7311 gcagattctagtagagtcacacacatactgcgtgtgaataccgcacagatgcgtaaggag 7370  
 2468 GCAGATTCTAGTAGAGTGCACCATATGCGTGTGAATACCGCACAGATGCGTAAGGAG 2527  
 QY 7371 aaataaccgcatcaggaatgtgtaacgttaataattgttaaaatcgcgttaaaattt 7430  
 2528 AAATATACCGCATCAGGAATGTGAAGCGTTATATTTTGTAAATTTCCGTTAAATTTT 2587  
 QY 7431 tgttaataacactcaatttttaaccaataagcgcaaaatcgcgcaaaatccctataatca 7490  
 2588 TGTTAATACACTCAATTTTAAACCAATAGCGCAAAATCGCAAAATCCCTTATAATCA 2647  
 QY 7491 aaagaataagcagagataggtgtgagtgtgtctcagtttggaacagaagtcacatca 7550  
 2648 AAAGAATAGACCGAGATAGGCTGTGAGTGTGTTCCAGTTTGGAAACAGAGTCCACATATTA 2707  
 QY 7551 aagaacgtgagctcccaacgctcaaaagcgcaaaacgctctcagcgagcgtgcccacta 7610  
 2708 AAGAACGTGAGCTCCAAGCTCAAAAGGCGCAAAACCGTCTATCAGGGCGATGGCCACATA 2767  
 QY 7611 cgtgaacataaccctaatcaagcttttgggtgtcgaggtgcgcgaagaacactaaatcg 7670  
 2768 CGTGAACCATACCTAATCAAGTTTTGGGTGTGAGGTGCCGTAATAACACATTAATTCG 2827  
 QY 7671 aaccttaaggaagcccgagattagagcttgaagcggaagcgcgagcgtgagcgata 7730  
 2828 AACCTTAAGGAGCGCCGATTTAGAGCTTGACGGGGAAGCCGCGCAACGTGGCGAGA 2887

QY 7731 aaggaagggaagaaacgaagagcgagcgctlaagcgcttgcgaagttagcggtcag 7790  
 2888 AAGGAAGGGAAGAAACGAAGAGCGCGCTAGAGCGCTGCAAGTGTAGCGGTACG 2947  
 QY 7791 ctggcggttaaccacacaccgcgcgcttaatggtcgctcagagcggttcgcccac 7850  
 2948 CTGCGCGGTAAACCAACACACACCCCGCGCTTAATGGCCGCTACAGGGCGCT---CCAT 3003  
 QY 7851 tgcacatcaggtacgaacactgttggaaggcgatcggtgcgagcctctcgtatata 7910  
 3004 TCGCATTTAAGGCTGCGCAACTGTTGTGGAAGGCGCATGCTCGGGCCTTTCGTATTA 3063  
 QY 7911 cgccaagctgctgcagggggggggg 7935  
 3064 CGCCAGCTGCGCAAAAGGGGATGTG 3088  
 Db 3064 CGCCAGCTGCGCAAAAGGGGATGTG 3088  
 RESULT 12  
 CVPGEN32P  
 LOCUS CVPGEN32P 3197 bp DNA circular SYN 16-JUL-1999  
 DEFINITION Cloning vector pGEM-32f(+).  
 ACCESSION X65306  
 VERSION X65306.2 GI:5531232  
 KEYWORDS Cloning vector; multiple cloning site; phage fl region; promoter.  
 SOURCE Cloning vector pGEM-32f(+).  
 ORGANISM  
 Solomon, L.R., Massom, L.R. and Jarrett, H.W.  
 Enzymatic syntheses of DNA-silices using DNA polymerase  
 Anal. Biochem. 203 (1), 58-69 (1992)  
 92398067  
 2 (bases 1 to 3197)  
 Technical Services.  
 Direct Submission  
 Submitted (23-MAR-1992) Technical Services, Promega Corporation,  
 2800 Woods Hollow Road, Madison, WI 53711-5399, USA  
 JOURNAL  
 REFERENCE  
 1 (bases 1 to 3197)  
 Technical Services.  
 Direct Submission  
 Submitted (28-MAR-1993) Technical Services, Promega Corporation,  
 2800 Woods Hollow Road, Madison, WI 53711-5399, USA  
 JOURNAL  
 REMARK  
 4 (bases 1 to 3197)  
 Kenefick, K.  
 Direct Submission  
 Submitted (14-JUL-1999) Kenefick K., Technical Services, Promega  
 Corporation, 2800 Woods Hollow Road, Madison, WI 53711-5399, USA  
 On Jul 19, 1999 this sequence version replaced g1:58203.  
 See X65300-X65335 for related vector sequences  
 This vector can be obtained from Promega Corporation, Madison, WI.  
 Call one of the following numbers for order or technical  
 information:  
 Order or Technical 800-356-9526  
 In Wisconsin 800-356-9526  
 Outside U.S. 608-274-4330.  
 Location/Qualifiers  
 1. 3197  
 /organism="Cloning vector pGEM-32f(+)"  
 /db\_xref="taxon:90112"  
 /note="T7 transcription initiation site"  
 5..61  
 /note="multiple cloning sites"  
 67..86  
 /note="SP6 promoter"  
 69  
 /note="SP6 transcription initiation site"  
 94..323  
 /note="lac operon sequence"  
 complement(106..108)  
 /gene="IacZ"  
 misc\_feature  
 misc\_feature  
 misc\_feature







Db 1946 GGGGCGAAAACCTCTCAAGGATCTTACCGCTGTTGAGATCCAGTGGANGTAACCCACTCG 2005  
 QY 6831 Tgcaccacatgaltctcagacatcttactcttcacccagcgtcttcggtgagcaaaac 6890  
 Db 2006 TGCACCCCAAGTATGATGAGCATCTTTACTTTCATCTTCAACAGCGTTTCGGGTGAGCAAAAC 2065  
 QY 6891 aggaagcagcaaaatgcgcgcaaaaaggaataagggcgacacggaaatgtltgaatactcat 6950  
 Db 2066 AGGAAGGCAAAATGCGGCAAAAAGGGAATAGGGCGACACGGAATGTGTAATCTCAT 2125  
 QY 6951 acctctccctttcacaatcttaagacatctatcagaagttatctcacaagaagata 7010  
 Db 2126 ACTCTTCTTTTCAATTTATTTAGACATTTATCAGGGTTATTTGTCTCATAGCGGATA 2185  
 QY 7011 cataattgaatglatltagaanaataacaataaggggttccgcgcacattcccgaaa 7070  
 Db 2186 CATATTGAATGTATTATGAAAAATAAACAAATAGGGGTTCGGCGACATTTCCCGGAA 2245  
 QY 7071 agtgcacatgagtcgtctcagaacacataataatcatatgacataacctataaataagcg 7130  
 Db 2246 AGTCCACCTACGCTCTAAGAAACCATTTATCATGACATTAACCTATAAAAAATAGCGC 2305  
 QY 7131 tatcagaagggccttctgcgcgcgttctgcgtatgataaggtgaaaacctctgacat 7190  
 Db 2306 TATACAGAGGCCCTTTCGTCTCGCGGTTTGATGATGAGGTGAACCTCTGACACAT 2365  
 QY 7191 gcaagctcccgagagacggtcacagcttgcgtgaaagcgatgcggagagacagaacggcg 7250  
 Db 2366 GCAGCTCCCGGAGAGAGGTGACAGCTTGTCTGTAACGGATGCGGGAGAGACAAAGCCG 2425  
 QY 7251 tcagggcgagtcagcgaggtgtgtgcgggtgtcggggcttggttaactabgcgcatcaga 7310  
 Db 2426 TCAGGGCGCGCTCAGCGGGGTGTTGGCGGGTGTGCGGGCTGCTTAATGCGCGCATCAGA 2485  
 QY 7311 gcaagttgactgagagaggtgacacatatcggtgtgaataaccgcacagatggtgaagag 7370  
 Db 2486 GCAGATTGTACTGAGAGGTGACACATATCGGGTGTAAATACCGCACAGATGGGTAAAGGAG 2545  
 QY 7371 aaaaatacgcatacaggaatgtgaacgttlaatatctgttaaaatctcgcgtlaaatct 7430  
 Db 2546 AAATATACCGCATCAGGAATTTGTAGCGTTATATTGTTGTTAAATTCGCGTTAAATTTT 2605  
 QY 7431 tgttaaatcagctcaatttttaaccaataagcgcaaatcgcgcaaaatccctataatca 7490  
 Db 2606 TGTTAATATCAGCTCATTTTAAACCAATAGGCCGAATCGGCCAAATCCCTTATTAATCA 2665  
 QY 7491 aaagaatagacggagataggttgaggtgtgttccagtttgaaagaagagtcactatca 7550  
 Db 2666 AAGGAATAGACCGAGATAGGTTGAGTGTGTTCAGTTTGGACACAGAGTCCACTATTA 2725  
 QY 7551 aagaacgtlgaatcccaacgtlcaaaagcgcaaaacgcttatacagggcgatlgccacta 7610  
 Db 2726 AAGAACTGAGCTCAACAGTCAAAAGGGGGAATAACCGTCTATCAGGGGATGGCCACGA 2785  
 QY 7611 cgtgaacatcaaccttaactcaagtttttgggttcgaaggtgcgttaagacttaacg 7670  
 Db 2786 CGTAAACATCACTCATTAAGTTTGGGTTCGAGTGCCTGAAGGACATTAATCGG 2845  
 QY 7671 aaacctaaaggaagcccccgaattagaagcttgaacggagaaagccggagaaagctgcgaga 7730  
 Db 2846 AACCTTAAGGAGACCCCGGATTAGACTTGAACCTTGACGGGGAAGCCCGCAACCTGGCAGA 2905  
 QY 7731 aaggaaggaagaaagcgaagagcgagcgctgaagggcgcttgcgaagtgaagcgctcacg 7790  
 Db 2906 AAGGAAGGAAGAAAGCAAGAGAGCGGCGCTAGGGCGCTGCGCAAGTGAAGCGGTCAAG 2965  
 QY 7791 ctgcggtlaaccacacacccgcgcgttaatgcgcgcgtacagggcgctgcgcgcac 7850  
 Db 2966 CTGCGGCTTAACACACACCGCGCGCTTATGCGCGCTTAACAGGGCGCGT---CCAT 3021  
 QY 7851 tcgagcatcaggtacgaacatgtttgggaagggcgatcggtgcgggacctcttcgtatca 7910  
 Db 3022 TCGGCATTACAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCCTTTTCGCTATTA 3081

QY 7911 cgcagctggtcgcaggggggggg 7935  
 Db 3082 CCCAGAGCTGGCAAGGGGGATGTG 3106

RESULT 13  
 CVMEM12P  
 LOCUS CVMEM12P 3221 bp DNA circular SYN 05-JUL-1999  
 DEFINITION Cloning vector pGEM-11zf(+).  
 ACCESSION X65313  
 VERSION X65313.2 GI:5420374  
 KEYWORDS beta-lactamase; bla gene; cloning vector; lacZ gene; multiple cloning site; phase I1 region; promoter; pUC/M13 primer binding site.

SOURCE  
 ORGANISM Cloning vector pGEM-11zf(+).  
 Cloning vector pGEM-11zf(+).  
 artificial sequence; vectors.  
 REFERENCE 1 (bases 1 to 3221).  
 AUTHORS Technical Services.  
 TITLE Direct Submission  
 JOURNAL Submitted (23-MAR-1992) Technical Services, Promega Corporation, 2800 Woods Hollow Road, Madison, WI 53711-5399, USA

REMARK Revised by [4]  
 REFERENCE 2 (bases 1 to 3221)  
 AUTHORS Technical Services.  
 TITLE Direct Submission  
 JOURNAL Submitted (05-JUL-1999) Technical Services, Promega Corporation, 2800 Woods Hollow Road, Madison, WI 53711-5399, USA

COMMENT  
 On Jul 8, 1999 this sequence version replaced g1:58168.  
 See X65300-X65335 for related vector sequences  
 This vector can be obtained from Promega Corporation, Madison, WI.  
 Call one of the following numbers for order or technical information:  
 Order or Technical 800-356-9526  
 In Wisconsin 800-356-9526  
 Outside U.S. 608-274-4330.

#### FEATURES

1..3221 Location/Qualifiers

1..3221 /organism="Cloning vector pGEM-11zf(+)"

1..3221 /db\_xref="taxon:90119"

1..3221 /note="T7 RNA polymerase transcription initiation site"

10..85 /note="multiple cloning region"

91..110 /note="SP6 promoter"

93 /note="SP6 RNA polymerase transcription initiation site"

118..347 /note="lac operator"

128..144 /note="pUC/M13 reverse primer binding site"

complement(130..132) /gene="lacZ"

complement(130..132) /gene="lacZ"

/note="lacZ start codon"

152..168 /note="lac operator"

complement(1289..2149) /gene="bla"

complement(1289..2149) /gene="bla"

/codon\_start=1

/transl\_table=1

/product="Beta-lactamase"

/protein\_id="CAA6405.2"

/db\_xref="GI:5420375"

/translation="MSIOHPRVALIPFAFCBPFAHPETLVKVDARFQGLGARYG

IELDINSKGLTISREPRERPMSTKIVLCGLVLRIDAGDQLRRIHYSONDIYE

YSPVTEKHLDIGKTVRELCSAATISDNTRANLLLTITGGPKELTAFILNMGDHYTRL

DRWEPELNEAIPNDERDPTMPVAMATTLLRLKLLGLLTLASRQOLLIDMWEADKVA6PL



	misc_feature	LRSLPAGWTFADKSGAGERSRBIITAALEDKPSRTIVYITTGSAITMDERKQIA EIGASLILKHW		
	misc_feature	2586..3041 /note="phage fl region"		
	misc_feature	3042..3202 /note="lac operon"		
	misc_feature	3162..3178 /note="pUC/M13 forward primer binding site"		
	promoter	3205..3221 /note="T7 promoter"		
BASE COUNT	826 a	793 c	825 g	777 t
ORIGIN				
Query Match	35.4%	Score 2809.4	DB 56	Length 3221
Best Local Similarity	99.5%	Pred. No. 0		
Matches 2830	Conservative	0	Mismatches	11
			Indels	4
			Gaps	1
QY	5091	ccaagctgcaatgaatcatcgccaaacgcgcggggaagcggttgcgtattgagcgtc	5150	
DB	290	CCAGCTGCAATTAATTAATACGGCCAAACGGCGGGAGAGCGGTGTCCTATTGGCGCTC	349	
QY	5151	ttccgccttcctgcgcacactgactgctgtcgcctcgttcttgcgtgcgaggcgatgc	5210	
DB	350	TTCCGCTTCTCGCTCGCTACGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT	409	
QY	5211	agctactcaaaaggcggtaataacggtttaccacaatcgaagggaataacgcagaagaa	5270	
DB	410	AGCTACATCAAAAGCGGGTAAATACGGTTATCCACAGAAATCAGGGGATTAACGACGAA	469	
QY	5271	catgtgaacaaaagccagcaaaagccagaagaccgtataaaagcgctgtgctgcgtc	5330	
DB	470	CATGTGACCAAAAGCGCCAGCAAAAGCCACGAAACCGTAAAAAGCGCGCTGTGCGCTT	529	
QY	5331	tttcgatagagctccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc	5390	
DB	530	TTTTCCATAGGCGCTCGCGCCCCCTCGACGACATCACAAAATGACCGCTCAAGTCA	589	
QY	5391	gcgaaccgcgaacagactataaagaataccagcgctttcccccctggaagctccctgc	5450	
DB	590	GGGAACCCGACAGGACATATAAAGTATACAGCGCTTCCCTCGAAGCTCCCTCGTGG	649	
QY	5451	ctctcgttctccgaccctgcgcgttaaccggaataacctgcgccttctccctcgggaag	5510	
DB	650	CTCTCCTGTTCCGACCCCTCGCGCTTACCGGATACCTGTCGCCCTTTCCTTCGGGA	709	
QY	5511	cgtagcgctttctcaagtctcaacgcttagagttatctcaagttcgtgtagtgcgtctc	5570	
DB	710	CGTGGCGCTTCTCTCACTACACGCTGTAAGTATCTCAGTTGCGTGAAGTGGTTCGCTC	769	
QY	5571	caagctggcgctgtgacacgcaaccccccgcttcaagccgaacccgctbgscttaccgta	5630	
DB	770	CAAGCTGGGCGTGTGTGACACGAACCCCCGTTACGCCGACCGCTGCGCTTATCCGGTA	829	
QY	5631	ctatgccttgaagtcacaccgcgtaagaacagacttaacgcacatgcgacatgcagcga	5690	
DB	830	CTATGCTGTGAGTCAACACCCGGTAAAGACAGACTTATCGCACATGCGACGACGAC	889	
QY	5691	taacaggaattgcagagcgagatgtagtgagcggtgctacagaagtcttgaagtgtgccc	5750	
DB	890	TAAACGGATTGACCAACGAGGACTATGTGGGGGCTCAACAAGTTCCTTGAAGTGGTGG	949	
QY	5751	taacacaggtcacacataagaagacgataattgtagtctcgcgcctcgtcgtgaagcga	5810	
DB	950	TAACTACGCGCTACACTAATAAGAACAGTATTGTTGGTATCTGCGCTGTGTAAGCC	1009	
QY	5811	cttcggaataaagatctgtagctcttgaatccggaacaaacacaccgcttgaacggtg	5870	
DB	1010	CTTCGGAATAAAGATTGTAGCTCTTGATCCGGCAAAACAAACACCGCTGTGTAACG	1069	
QY	5871	tttttttcttgcgaagcagcagatagcgcgcaaaaaaagaatcctaagaagatccttc	5930	
DB	1070	TTTTTTTTTTTTCGCAAGCGACGAGATTACCCGCGAATAAAAGATCTCAAGTAAGATCC	1129	

QY	5931	gaccttctctacggggtctgaagctcagctgagcaagcaaacacglttaaggatcttggct	5930
Db	1130	GATCTTTTCATCGGGGGTGTGACGCTCAGTGGAAAGCAAACTACGTTAAGGGATTTTGGT	1189
QY	5991	catgagatctacaaaagagatcttcacatgatctccctttaattaaatgaaagtcttaa	6050
Db	1190	CATGAGATTATCAAAAAGAGATCTTTCACCTTGATCCTTTTAATTTAAATGAAAGTTTAA	1249
QY	6051	atcaactaaagatataatagtaagaaacttgcctcagttacagttaccatgcttaacatgta	6110
Db	1250	ATCATCTTAAAGGTATTAATATGATTAACCTTGCTGTGACAGTTTACCAATGCTTAAATCAGTA	1309
QY	6111	ggacacctctccacagagctgcgtcatcttctgtatctcaatcaatgcttgctgaatcccgctgt	6170
Db	1310	GGCACTATCTTCACAGCATCTGTCTATTTCGTATCATCAATAGTGGCTGTGACTCCCGCTGT	1369
QY	6171	gtagaataactacgaatacgggagaggtcttacatcttgcccccagtgctcgaatatalaccg	6230
Db	1370	GTAGATATCTACGATACGATACGGGGGCTTACATCTGGCCCCAGTGGCTGAATGATACGGG	1429
QY	6231	agaccacagctctaaccggtctccagattatcagcaataaacccagccagccgaagggccga	6290
Db	1430	AGACCCACAGCTCACCGGCTCCAGATTTATTCAGCAATTAACACAGCCAGCCGGAAGGGCCGA	1489
QY	6291	ggcgaagaagtgtctctcgaacttataccgctctcatcagatctatataattgtctgcggga	6350
Db	1490	GCGCAGAAAGTGGTCTCTCAACTTTATTCGCCCTTCATTCGACTATTAATTTGTTGCCGGGA	1549
QY	6351	agctagaagtaagtagtctgcgcagttaaagttctgcgaacagttgttgcattgtctacag	6410
Db	1550	AGCTAGAGTAAAGTGTTCGCGCAGATTAATATTGTTGCCCAAGTGTGTGCATGTCTACAG	1609
QY	6411	catcgttgctgcgaagctcgtcgttttgtagtgcctatccagctcccggtctcccaagctc	6470
Db	1610	CATGCTGGTGTGACAGCTCGTGTGTTGGTATAGGCTTATTCAGCTCCGGTTCCTCAACAGATC	1668
QY	6471	aaggcgagttacacgatcccccacatgttctgtaaaaaaagcggtttagctccctcgtccctc	6530
Db	1670	AAGCGAGTTTACATGATTCGCCCATGTTGTGCAAAAAAGGGTTAGCTCTCTTGGTCTTC	1729
QY	6531	gacgtgtgtcagaagaagtaagttggccgagtggtatataccatcagtggtatagcagacgtca	6590
Db	1730	GATCGTGTCTGAAAGTAAAGTTGGCCGAGAGTATACATCATGATGTTATGGCAGCACTCCA	1789
QY	6591	taattctcttaactctcatgcacatccgtaagaatgcttctctgacatggtgaatactaac	6650
Db	1790	TAAATTCCTTACTGTCAATGGCATTCGGTAAGATCCTTTTCTGTGACGTGGTGAATACTAAC	1849
QY	6651	caagtcaltctgagaatagtgtaatgcgcgcagcttgcctctgcgcgcgctcaatcag	6710
Db	1850	CAAGTCATTTCTGAAATATAGTATATGGGGCAGCCAGATTTCTCTTGGCCGCGCTCAATACG	1909
QY	6711	ggataataacccggccacaaatagcagaactttaaagtgccatcatattgtaaaaagctttctc	6770
Db	1910	GGATATATACCGCGCCACATATGCGAAGACTTTAAAGTGGCTCATATTGGAAAAAGTTCTTC	1969
QY	6771	ggggcgaaaaactctccaaagatactctacgcgtgtttagagatccagttcgaatgtaaccacag	6830
Db	1970	GGGGCGAAAAACTCTCAAGGATCTTACCGCGTTGAGATCCAGTTCGATGTAAACCATCTCG	2029
QY	6831	tgcaccacaactgatactctcagcatccttctactctcaacagcgcttctcgtgtgagcaaaac	6890
Db	2030	TGCACCCAAACGTATCTTCACACTCTTTTACTTTTCACACACGTTTTCTGGGTGCGCAAAAAC	2089
QY	6891	aggaagaacaaatgycgcgaaaaaaggaataaaggcgagcaaggaagatcttgataactcat	6950
Db	2090	AGGAAGGCAAAATGCCCCCAAAAAGGAATTAAGGCGGACAGCGGAATGTGAATCTCTAT	2149
QY	6951	acctctccctttccaatatattgaagcalttaacaggttatcgaaggtatgctcactaagcagata	7010
Db	2150	ACTCTTCCTTTTCAATATATTAATGAAGCAATTTATCAGGGTTATTTGCTCATACAGCGGATTA	2209



Qy 7011 catattgaaatglatcttaagaanaaatacaacaaatagagggtccgcgcacattccccgaaa 7070  
|||||  
Db 2210 CATATTGAAATGATATTGTAAGAAAATAAACAATAGGGGTCCGCCACATTTCCCGGAAA 2269  
Qy 7071 agtgcaccactagcgtctagaacacattatctatctgacattacactataaataatagagg 7130  
|||||  
Db 2270 AGTGCACCTACGCTATGAGAAACCATATTATATCATGACATTACCTATATAAATAAGGGG 2329  
Qy 7131 tacaagagggccctctcgtccgcgtctcgtgatgacagg tgaanaaccttgacacat 7190  
|||||  
Db 2330 TATACAGAGGGCCCTTTCGCTCGCGCGTTTCGGTCATGACGGGTGAANAACCTTGACACAT 2389  
Qy 7191 gcagctcccgagagacggttcaacagcttctctgtaagcggatgcgggagcagacaaagccg 7250  
|||||  
Db 2390 GCACCTCCCGGAGACGGTCACAGAGCTTGCTGTAAACGGATGCGGGAGACAGCAAGCCCG 2449  
Qy 7251 tcaagggcgcgctcagcgaggttggtggcggtgctgggggtctgaactaagcgacatcaga 7310  
|||||  
Db 2450 TCAGGGCGCGTCAGCGGGGTGGGGGTGGGGGTGGGGTGAACCTATGCGGCATCAGA 2509  
Qy 7311 gcagattgactgagagtgcaaccatactgcgtgtgtaataacgcgcacagatgctgaagag 7370  
|||||  
Db 2510 GCATATTGCTACTGAGAGTGCACACATATCGGTGTGAATACCGCACAGATGGCTAAGAG 2569  
Qy 7371 aaaaataccgcatcaggaaattgtlaaacgtlaaattctgtaaaatlccgcgtlaaattt 7430  
|||||  
Db 2570 AAAATACCGCATCAGGAATTTGTAAAGCTTAATTTGTTAAATTTCCGCTTAAATTTT 2629  
Qy 7431 tgttaaatcaactatatttttaacaaatagggcgaatcggaaataccctataatca 7490  
|||||  
Db 2630 TGTTAATCACACTCTATTTTAAACCAATAGCGCCGAATCGGCAANAATCCCTTATTAATCA 2689  
Qy 7491 aaagaatagacccgagatalagggttgagtgltgltccagttctgnaaagaagtcacatata 7550  
|||||  
Db 2690 AAAGAATAGACCGAGATAGGGTGTGAGTGTGTTCCAGTTTGGAAACAGAGTCACATATTA 2749  
Qy 7551 aagaagctgagctcccaacgltcaaaaggcgaaaaacggtctatcaaggcgatgcccacta 7610  
|||||  
Db 2750 AAGAACGTGGACTCCMAAGTCAAAAGGGGCAAAACCGTCTATCAGGGCCATGGCCACATA 2809  
Qy 7611 cgtlaaacatacccttaataagtttctgaggctcgaagtgccgtlaaaacataatcgg 7670  
|||||  
Db 2810 CGTAAACCATATACCTTAATCAAGTTTGTGGGGTGCAGGTGCCGTAAACACATAAATCCG 2869  
Qy 7671 aacctaaaggagcccccgatlltagagcttgaacggggaaaagccggcgagacgttgcgaga 7730  
|||||  
Db 2870 AACCTAAAGGAGGAGCCCCGATTTAGAGCTTGACGGGGAACCGCGCAACCTGGGGAGA 2929  
Qy 7731 aaggaagggaagaaagcgaaaggagcggtcgttagaggcgcttggaagtgtagcggtcacg 7790  
|||||  
Db 2930 AAGGAAGGGAAGAAAGCAAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTACCGGTCAAG 2989  
Qy 7791 ctggcggttaaacacacacaccccgcgcttaatgagcggtcgaaggcggtcgcgccat 7850  
|||||  
Db 2990 CTGGCGGTAAACACACACACCGCGCTTANTGCGCCCTACAGGGCCGCT---CCAT 3045  
Qy 7851 tccgcaattcagggtctacgaaactgtctggaagggcgatcggtgcgggacctcttcgcatata 7910  
|||||  
Db 3046 TCGGCATTCAGAGCTGCGCAACGTTGGGAAGGGCGATCGGTGCGGCGCTTCCTCGCATTA 3105  
Qy 7911 cgccaagctggtctgcaagggtggggg 7935  
|||||  
Db 3106 CGCCAGTGGCGCAAGGGGGATGTG 3130

RESULT 14  
A93963 4118 bp DNA PAT 22-JAN-2000  
LOCUS A93963  
DEFINITION Sequence 17 from Patent MO9718322.  
ACCESSION A93963  
VERSION A93963.1 GI:6742064  
KEYWORDS  
SOURCE unidentified.

ORGANISM unidentified  
unclassified.  
REFERENCE 1 (bases 1 to 4118)  
AUTHORS Dahm,M.W.  
TITLE METHOD OF QUANTIFYING TUMOUR CELLS IN A BODY FLUID AND A SUITABLE  
TEST KIT  
JOURNAL Patent: WO 9718322-A 17 22-MAY-1997;  
DAHM MICHAEL W (DE)  
FEATURES  
source Location/Qualifiers  
1..4118  
/organism="unidentified"  
/db\_xref="taxon:32644"  
BASE COUNT 991 a 1032 c 1094 g 1001 t  
ORIGIN  
Query Match 35.2%; Score 2794.2; DB 9; Length 4118;  
Best Local Similarity 99.3%; Pred. No. 0;  
Matches 2828; Conservative 0; Mismatches 13; Indels 6; Gaps 2;  
Qy 5091 ccagctgcatlaatgaatcgcgaacgcgcgggagagcggttgcgtatctggcgctc 5150  
|||||  
Db 1185 CCAGCTGCATTATATGATATCGCCACAGCGGGGAGAGCGGTTTGGTATGGGGCTC 1244  
Qy 5151 ttccgcttccctgcgtcactgactgcgtgcgtcgggttcggtcgtgcggagcgatc 5210  
|||||  
Db 1245 TTCCCTCTCTCTGCTACAGTACCTCGCTCGCTCGCTCGCTCGCTCGCGGAGCGGTATC 1304  
Qy 5211 agctcaactaaaggcggttaatacgggttaccagaatacagggttaacgagaagaa 5270  
|||||  
Db 1305 ACCTCACTAAAGGGGATATAGGTTATCCACAGATCAAGGGATACGAGAAAGAA 1364  
Qy 5271 catgtgagcaaaaagccagcaaaaagccaggaacggtlaaaaggccggttgcgtgcgt 5330  
|||||  
Db 1365 CATGTGAGCAAAAGCCAGCAAAAGCCAGGAACCGTAAAGGGCGGCTTGCGGCTT 1424  
Qy 5331 ttccaataggtccgccccctgacgagcatcaaaaaatcgaagctaaagtcaagggtg 5390  
|||||  
Db 1425 TTTTCATAGGCTCCGCCCCCTGAGGACATCACAAAATGAGCGCTCAAGTCAGAGGTG 1484  
Qy 5391 gggaaacccgagacactataaagataccagcggttcccccctggaagctccctcgtgcg 5450  
|||||  
Db 1485 GCGAAACCCGACAGAGACTATTAAGTATACCAGGCGTTTCCCTTGGAAAGCTCCCTCGTGG 1544  
Qy 5451 ctctcgttctccgaacccgcgtcaccggaataccgtctgcgtcttctccctcgggaag 5510  
|||||  
Db 1545 CTCCTCGTTCGACCGCCGCGCTTACCGGATACCTGTCGCCCTTTCCTTCGGGAG 1604  
Qy 5511 cgtgcgcttctcgaatgctcaagctgtagatatacgaagtcggttgtagtgcgtc 5570  
|||||  
Db 1605 CTTGCGCTTCTCTATACCTACCGCTGTAGGTATCTCAGTTGCGGTGAAGTCTCCCTC 1664  
Qy 5571 caagcttgagcggtgtgacggaaccccgcttcagccgaacccgcgtccctataccgtaa 5630  
|||||  
Db 1665 CAAGCTGGGCTGTGTGACAGAACCCCGCTTACGCCGACCGCTGCGCTTATCCGGTAA 1724  
Qy 5631 ctatcgcttgaagtcnaacccggtlaagaacagacttaccgacactggaagcagccactg 5690  
|||||  
Db 1725 CTATCGTGTGAGTCCAACCCGGTAAAGACAGACTTATCGCACATGGAGCAGCACATGG 1784  
Qy 5691 taacaggaattagcagagcgaaggtatgtagcggtgtcagagatctctgaaggtgtggc 5750  
|||||  
Db 1785 TAACAGGATTTAGCAGAGGAGGTATGTAGGGGCTTCAAGAGTCTTGAAGTGTGGGCC 1844  
Qy 5751 taactagcgtacactagaagaagataattgtaatactgcgtcgtcgtgaagcaggttac 5810  
|||||  
Db 1845 TAACTACGGCTACACTAAGAGGACAGTATTTGTATCTGCGCTTGCTGAAGCCAGTTAC 1904  
Qy 5811 cttcgnaaaaaggttggtagctcttga tccggaacaaacacacccgttggtagcgttgg 5870  
|||||  
Db 1905 CTTGGAAAGAGTTGGTAGCTTTGATCCGGCAACAAACACACCGCTGTTACGGGTGG 1964  
Qy 5871 ttlttctgttgcaagcgacagattaagcgcgcaaaaaaaggatctcaagaagatccctt 5930



|||||  
Db 1965 TTTTGTGTCAGACGAGATTACGCGAGAAAAAGATCTCAGAGATCTTT 2024  
Qy 5931 gatctttctcggggtctgacgtcagtggaagaaaacccaagttaagatcttggt 5930  
Db 2025 GATCTTTCAGGGGTCGTGAGCTCACTGAGGAGAAACCTACAGTTAAGGATTTTGGT 2084  
Qy 5991 catgagatatacaaaaagagatctcacctacatcctcttaaatcaaaatgaatttaa 6050  
Db 2085 CATGAGATTATCAAAAAGATCTTACCTAGATCTTTAAATTAAATGAATTTTAA 2144  
Qy 6051 atcaatctaaagiatatataagtaaaacttggctcgcagagttacaaatgctaaatcag 6110  
Db 2145 ATCATCTAAAGTATATATAGTAAACTGGCTGACAGATTACCAATGCTTAATCAGTGA 2204  
Qy 6111 ggcacctatctcagcgatctgtctatcttgctcaatcagtgctccgaccccgctcgt 6170  
Db 2205 GGCACCTATCTCACCGAGTCTGCTATTTTCATTCATCATTAGTTCCGACCTCCCGTCTG 2284  
Qy 6171 gtagataactacgatacgggagggcttaacatctggcccgagtgctgaatgataccgg 6230  
Db 2265 GTAGATTAACCTAGATACGAGGAGGCTTACCATCTGGCCCGCTGCTGCAATGATACCGCG 2324  
Qy 6231 agaccacgctcaccgggtccagattatcaagcaataaacagccagccggaaaggcgga 6290  
Db 2325 AGACCCACGCTCACCGGCTCCAGATTATCAGCAATTAACACGACGACCGGAGGCGCGA 2384  
Qy 6291 ggcgaagaagtgctcgtcaacttatccgctccatccagctctataatgcttgccggga 6350  
Db 2385 GCGGAGAGTGCTCTCACTTATCCGCTCCATCCAGCTATTAATTTGTTCCCGGGA 2444  
Qy 6351 agctagaagtaagtgctcgcagtgtaagtctgcgcaacgtctgcatctgtctacag 6410  
Db 2445 AGCTAGAGTAGTAGTGTCCGCGAGTTATAGTTGCGCAAGCTTGTGCCATGTCTAACAG 2504  
Qy 6411 catcgtgtgtcagcgtcgtcgttggctatggtctcaatcgaagtcgggtcccaagcgtc 6470  
Db 2505 CATGCTGTGTCAACGCTCGTGTGTGATAGCTTCAATCAGCTCCGCTCCCAACGATC 2564  
Qy 6471 aagcgaggttaacgatcccccagctgtgtgcaaaaaggggtgagctcccttgctctcc 6530  
Db 2565 AAGCGAGTTATCATGATCCCATGTTGTGCAAAAAGCGGTAGCTCTTCGGTCTCTCC 2624  
Qy 6531 gatcgtgtcagaagtaagtggtgcgcagtgctatcaactcagtgctatgagcagactgca 6590  
Db 2625 GATGTTGTGAGAGTAGTGTGGCGCAGTGTATCATCTATGTTATGGAGCAGCTGCA 2684  
Qy 6591 taattctctactgtcatgcccactcgtlaagatgctcttctcgtgagcgtgtgagtaacctaac 6650  
Db 2685 TAAATCTCTTACTGTCAATGCCATCCGTAAGATGCTTTCTGTGACTGCTGATCTCAAC 2744  
Qy 6651 caaatcatctcgaagaatgaagtatgtagcggaacggaagtgtgctcccgcgctaaatcag 6710  
Db 2745 CAATCTATCTTGAAATATGATATGCGCGACCGAGTTGCTCTTGCCCGGGGTAAATACG 2804  
Qy 6711 ggaataatacgcgcacatagcagaactttaaagtgctcatcaatctggaaaacgtctctc 6770  
Db 2805 GGATTAATACCGCGCCACATGACAGACTTTAAAGTGTCTATCTTGGAAAAAGCTTTCTTC 2864  
Qy 6771 ggggagaaaactctcaagatcttaacgctgtgtgagatccaagtctgaatgaaccactcg 6830  
Db 2865 GGGGCGAAAACTCTCAAGGATCTTACCGCTGTGAGATCCAGTTGATGAACCACTCTCG 2924  
Qy 6831 tgcacccaactgactctcaacatcttctacttcaacagaggtctcgtggaagaaaac 6890  
Db 2925 TGACCCCACTGATCTTCAGCATCTTTTACTTTTACCAAGCGTTTCTGGGTGAGAAAAAC 2984  
Qy 6891 aggaaggcaaaatgcccgaaaaaaagggaatlaaggcgacacggaatgctgaatactcat 6950  
Db 2985 AGGAAGGCAAAATGCCGCAAAAAAGGAATTAAGGCGACACGGAATGTTGAATACCAT 3044  
Qy 6951 actctccttttcaatatatgaagcaattatcaaggttatctgtctcagcgagata 7010  
|||||

Db 3045 ACTCTCTCTTTTCATATATTATTAAGCATTTATCAGGGTATTTGTCATGAGCGGATA 3104  
Qy 7011 catattgaagtatattagaanaaataaacaatatggggttcgcgcacatttcccgaaa 7070  
Db 3105 CATATTTGAATGTATTTAGAAAAATAACAATATAGGGGTTCGCGCACATTTTCCCGAAA 3164  
Qy 7071 agtccacactcagctcctaagaacacatataatcatatgacattaaactataaataagcg 7130  
Db 3165 AGTCCACACTCAGCTCTTAAGAAACCATTTATCATGATCATTAACCTATAAATTAAGCG 3224  
Qy 7131 tatacgaaggcccttcgctcgcggttcctggtatgacgtgtgaaaaactctgcacat 7190  
Db 3225 TATCAGAGGCGCTTTCGTCTCGCGCTTTCGTGATGACGCGTGAACCTCTGACACAT 3284  
Qy 7191 gcaagctccggagagcgttcaacagctgtctgtaagcggaatgcccggagccagacagccg 7250  
Db 3285 GCAGCTCCCGAGACGGGTCAACAGCTTGTGTAAAGCGATGCGCGGAGCAACAAGCCG 3344  
Qy 7251 tcaaggcgctcagcggtgtgtgaggtgtgtcggggtggttaactatgagcatcaga 7310  
Db 3345 TCAGGGGCGCTCACGGGCTGTGGCGGCTGTGGGGCTGGCTTAACATATGGCGCATCAGA 3404  
Qy 7311 gcaagattgtaactgagtgcaacatagcggtgtgtaataacgcagacagatgctlaaggag 7370  
Db 3405 GCAGATTGTACTGAGAGTGCACCATATGCGGTGAAATACCGCACAGATGCTTAAGAG 3464  
Qy 7371 aaaaaccgacatcg--gaattgtaaggttaatttttgtaaaattgcggttaatt 7428  
Db 3465 AAAATACCGCATCAGGCGAAATTTGTAACGTTAATTTTGTAAATTCGCGTTAATA 3524  
Qy 7429 ttgttaaatcagctcatcttctaacaataagcggaatcgcgaataatccctataat 7488  
Db 3525 TTTCTTAATCAGCTCATTTTTTTAACCAATAGCGCAATGCGAAATCCCTTATTAAT 3584  
Qy 7489 caaagaatagaaccgagataggggtgaggtgtgtccagtttggacaagaagttccactat 7548  
Db 3585 CAAAGGATATGACCCGAGATAGGGTGTGAGTGTTCACAGTTTGGAACAAGATCCACTAT 3644  
Qy 7549 taaagaacgtgagactccaagctcaaaaggcgaaaacggtctatcaaggcgatgcccac 7608  
Db 3645 TAAAGAACTGTGACTCCACACTCAAGGCGAAAGCGTATCAAGGCGATGCGCCAC 3704  
Qy 7609 taactgaacatcaaccccttaactcaagttcttggtgtcgaagtggtgcgttaaacataatc 7668  
Db 3705 TAGCTGAACCATACCCCAATCAAGTTTTCGCGTGAAGTGCCTTAATATC 3764  
Qy 7669 ggaaccctlaaaggagcccccagatttagagcttgaacggggaagccggcgaaactgtgcga 7728  
Db 3765 GGAACCTTAAGGGAGGCCCGCGATTGTAGACTTGACGGGGAAGCCCGCAACGTTGGCGA 3824  
Qy 7729 gaaagggaagggaagaagggaagggaagcggtctgaaggcgtggaagtgtagcggtca 7788  
Db 3825 GAAAGGAAGGAAGAAAGCAAGAAAGGAGCGGCTTAAGGCGCTGAGCAAGTGTACCGGTCA 3884  
Qy 7789 cgctgcgcgtlaaacacacacacccgcgcgttlaatggtgcgcgttacaaggtgcgtgcgc 7848  
Db 3885 CGCTGCGCTTAACACACACACACCGCGCTTATGCGCGGCTTAACAGAGGGCGGT---CC 3940  
Qy 7849 attcgccattcgaagctacgcgaactgttgggaaggcgcatcgtgtgcggtctctgctat 7908  
Db 3941 ATTTCGCTATTCAGGCTGCGCAACTGTGGGAAGGCGCATCGTGCGGGCTCTTCCTAT 4000  
Qy 7909 taagcagcgtgcgtcagggggggggg 7935  
Db 4001 TAGCGCAGCTGCGAAAGGGGATGTG 4027  
|||||

RESULT 15  
A93964 LOCUS A93964 4118 bp DNA  
DEFINITION Sequence 18 from Patent WO9718322.  
ACCESSION A93964  
VERSION A93964.1 GI:6742065



[illegible]

QY	587.1	ttt tg tgc aac gac gac gt aac gc gc g a a a a a a a g at c c aac ga g at c c t t	5930
Db	1965	TTTTTTT g TT G C A G C A G C A G G AT AC G C C A A A A A A A G G AT C C A G A A G AT C TT	2024
QY	593.1	g at c t t c t c a c g g g t c t c a g c t c a g t g a a c g a a a c a c t c a g t a a g g t t g t	5990
Db	2025	G AT C T T T T T C A C G G G G T C T A C G T C A G T G A A C G A A A A A C T C A G T T A G G AT T T G T	2084
QY	599.1	c at g a g a t t a c a a a a g a t t c t c a c t a g a t c c t t t a a t t a a a t g a g t t t a	6050
Db	2085	C AT G A G AT T A T C A A A A A G G AT C T T C A C T A G AT C C T T T A A T T A A A T T A A G A G T T T T A	2144
QY	605.1	a t c a a c t a a a g t a t a t a t a g t a a a c t g t c t g a c a g t t e c a a g t g t c a t g a	6140
Db	2145	A T C A T C A T A A G T A T A T A G T A A A C T T G T G A C A G T T A C A A T G T T A T A C A G T A	2204
QY	611.1	g gc a c t a c t c a g a g a t c t c t a t t c g t c a t c a t a g t g c t g a c c c g t c g t	6170
Db	2205	G G C A C AT C AT C A C G AT C G AT T T G T T C G T T C AT C A A G T T G G C T G AT C C C G T G T	2264
QY	617.1	g t a a t a a c t a c a t a a g g a g g c t t a c a t c t g c c c c a g t g t c g a t a t a c c g	6230
Db	2265	G T A AT A T A C T A C G AT A C G G A G G G C T T A C AT C T G C C C C A G G T G C A T A T A C C G	2324
QY	623.1	a g a c c a c g t c a c c g g c t c a g a t t a t c a g a a t a a c a a c a g c a g c c g a a g g c c g a	6290
Db	2325	A G A C C A C G G C A C C G G C T C A G AT T T A T A C A G A T A A A C A G C A G C G G A G G G C G A	2384
QY	629.1	g c g c a a a g t g t c t c t g a a c t t a t c c g c c t c a t c a g t c t a t t a t t g t c c g g a	6350
Db	2385	G C G C A G A G G G T C T C A C A T T T A T C G G T C C A T C A G T C A T A T T A T T G T T G C G G G A	2444
QY	635.1	a g c a g a t a g t a g t c g c a g t t a a t g t g c a a g t g t c g c a g t g t c g c a g t c a g	6410
Db	2445	A G C T A G A G T A A G T T G G C A T T A T A G T T T G G C A A C G T T G G C A T T G C A T T G C T A C A G	2504
QY	641.1	c a t c g t g t g t c a g c t c g t c g t t g t g t a g t c a t c a t c a g t c c g t c c c a g a t c	6470
Db	2505	C AT G T G T G T C A G C T C G T C G T T G T A T G G T T C AT T C A C T C G G T T C C A A C A G AT C	2564
QY	647.1	a a g c a g t t a c a t g a t c c c c a t g t c g c a a a a a a g g t a g t c t c t c g t c c c c	6530
Db	2565	A A G C A G A G T T A C AT G AT C C C C A T G T T G T G T C A A A A A A G G G T T A G C T T C G G T C T C	2624
QY	653.1	g a t c g t g t c a a a g t a g t t g c c g a g t g t a t c a c t a c a t g t a t g c a g c a t g c a	6590
Db	2625	G AT C G T T G T C A A A G T T G G C C A G G T T A T C AT C AT G T T A T A G C A G C A C T G C A	2684
QY	659.1	t a a t c c t a c t g t a t g t c a t c c g t a a g a t g c t t t c t g a c t g t g a t a c t a a c	6650
Db	2685	T A A T T C T T A C T G T C A T G G C A T C G T A G A G A G A G C T T T T G T G A C G G G A G A T A C T A C	2744
QY	665.1	c a a t c t c t g a a t a g t a t g c g c a c c a g t g t c t c t g c c g c g t c a a t g c a t g c	6710
Db	2745	C A A T C A T T C T G A N A T A G T A T G G G G C A C G A G A T T C T T G C C G G C T C A T A T A G C	2804
QY	671.1	g g a t a t a c c g c c a t a g c a g a c t t a a a g t c a t c a t c a t t g a a a c g t t c t c	6770
Db	2805	G G A T A A T A T C G C C C A C AT A G C A G A C T T T A A A G G T G C AT C AT T G G A A A A C G T T T C	2884
QY	677.1	g g g g c g a a a c t c a a g t c t a c c g t g t t g a t c c a t t g a t g t a a c c a c g	6830
Db	2865	G G G G C G A A A C	







**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 18, 2001, 04:30:13 ; Search time 11937.4 Seconds  
(without alignments)  
3497.675 Million cell updates/sec

Title: US-09-689-430-1\_COPY\_419\_4835

Perfect score: 4417

Sequence: 1 catgcaatagagctctcca.....agcacagagacctctactga 4417

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 10228115 segs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

```
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est25:*
26: gb_est26:*
27: gb_est27:*
28: gb_est28:*
29: gb_est29:*
30: gb_est30:*
31: gb_est31:*
32: gb_est32:*
33: gb_est33:*
34: gb_est34:*
35: gb_est35:*
36: gb_est36:*
37: gb_est37:*
38: gb_est38:*
39: gb_est39:*
40: gb_est40:*
41: gb_est41:*
42: gb_est42:*
43: gb_est43:*
44: em_esthum10:*
45: em_esthum11:*
46: em_esthum12:*
47: em_esthum13:*
48: em_esthum14:*
49: em_esthum15:*
50: em_esthum16:*
51: em_esthum17:*
52: em_esthum18:*
53: em_esthum19:*
54: em_esthum20:*
55: em_esthum21:*
56: em_esthum22:*
57: em_esthum23:*
58: em_esthum24:*
59: em_esthum25:*
60: em_esthum26:*
61: em_esthum27:*
62: em_esthum28:*
63: em_esthum29:*
64: em_esthum30:*
65: em_esthum31:*
66: em_esthum32:*
67: em_esthum33:*
68: em_esthum34:*
69: em_esthum35:*
70: em_esthum36:*
71: em_esthum37:*
72: em_esthum38:*
73: em_esthum39:*
74: em_esthum40:*
75: em_esthum41:*
76: em_esthum42:*
77: em_esthum43:*
78: em_esthum44:*
79: em_esthum45:*
80: em_esthum46:*
81: em_esthum47:*
82: em_esthum48:*
83: em_esthum49:*
84: em_esthum50:*
85: em_esthum51:*
86: em_esthum52:*
87: em_esthum53:*
88: em_esthum54:*
89: em_esthum55:*
90: em_esthum56:*
91: em_esthum57:*
92: em_esthum58:*
93: em_esthum59:*
94: em_esthum60:*
95: em_esthum61:*
96: em_esthum62:*
97: em_esthum63:*
98: em_esthum64:*
99: em_esthum65:*
100: em_esthum66:*
101: em_esthum67:*
102: em_esthum68:*
103: em_esthum69:*
104: em_esthum70:*
105: em_esthum71:*
106: em_esthum72:*
107: em_esthum73:*
108: em_esthum74:*
109: em_esthum75:*
110: em_esthum76:*
111: em_esthum77:*
112: em_esthum78:*
113: em_esthum79:*
114: em_esthum80:*
115: em_esthum81:*
116: em_esthum82:*
```



117: gb\_est48:\*  
118: gb\_est49:\*  
119: gb\_est50:\*  
120: gb\_est51:\*  
121: gb\_est52:\*  
122: gb\_est53:\*  
123: gb\_est54:\*  
124: gb\_est55:\*  
125: gb\_est56:\*  
126: gb\_est57:\*  
127: gb\_est58:\*  
128: gb\_est59:\*  
129: gb\_est60:\*  
130: gb\_est61:\*  
131: gb\_est62:\*  
132: gb\_est63:\*  
133: gb\_est64:\*  
134: gb\_est65:\*  
135: gb\_est66:\*  
136: gb\_est67:\*  
137: gb\_est68:\*  
138: gb\_est69:\*  
139: gb\_est70:\*  
140: gb\_est71:\*  
141: gb\_est72:\*  
142: gb\_est73:\*  
143: gb\_est74:\*  
144: gb\_est75:\*  
145: gb\_est76:\*  
146: gb\_est77:\*  
147: gb\_est78:\*  
148: gb\_est79:\*  
149: gb\_est80:\*  
150: gb\_est81:\*  
151: gb\_est82:\*  
152: gb\_est83:\*  
153: gb\_est84:\*  
154: gb\_est85:\*  
155: gb\_est86:\*  
156: gb\_est87:\*  
157: gb\_est88:\*  
158: gb\_est89:\*  
159: gb\_est90:\*  
160: gb\_est91:\*  
161: gb\_est92:\*  
162: gb\_est93:\*  
163: gb\_est94:\*  
164: gb\_est95:\*  
165: gb\_est96:\*  
166: gb\_est97:\*  
167: gb\_est98:\*  
168: gb\_est99:\*  
169: gb\_est100:\*  
170: gb\_est101:\*  
171: gb\_est102:\*  
172: gb\_est103:\*  
173: gb\_est104:\*  
174: gb\_est105:\*  
175: gb\_est106:\*  
176: gb\_est107:\*  
177: gb\_est108:\*  
178: gb\_est109:\*  
179: gb\_est110:\*  
180: gb\_est111:\*  
181: gb\_est112:\*  
182: gb\_est113:\*  
183: gb\_est114:\*  
184: gb\_est115:\*  
185: gb\_est116:\*  
186: gb\_est117:\*  
187: gb\_est118:\*  
188: gb\_est119:\*  
189: gb\_est120:\*

190: gb\_est121:\*  
191: gb\_est122:\*  
192: gb\_est123:\*  
193: gb\_est124:\*  
194: gb\_est125:\*  
195: gb\_est126:\*  
196: gb\_est127:\*  
197: gb\_est128:\*  
198: gb\_est129:\*  
199: gb\_est130:\*  
200: gb\_est131:\*  
201: gb\_est132:\*  
202: gb\_est133:\*  
203: gb\_est134:\*  
204: gb\_est135:\*  
205: gb\_est136:\*  
206: gb\_est137:\*  
207: gb\_est138:\*  
208: gb\_est139:\*  
209: gb\_est140:\*  
210: gb\_est141:\*  
211: gb\_est142:\*  
212: gb\_est143:\*  
213: gb\_est144:\*  
214: gb\_est145:\*  
215: gb\_est146:\*  
216: gb\_est147:\*  
217: gb\_est148:\*  
218: gb\_est149:\*  
219: gb\_est150:\*  
220: gb\_est151:\*  
221: gb\_est152:\*  
222: gb\_est153:\*  
223: gb\_est154:\*  
224: gb\_est155:\*  
225: gb\_est156:\*  
226: gb\_est157:\*  
227: gb\_est158:\*  
228: gb\_est159:\*  
229: gb\_est160:\*  
230: gb\_est161:\*  
231: gb\_est162:\*  
232: gb\_est163:\*  
233: gb\_est164:\*  
234: gb\_est165:\*  
235: gb\_est166:\*  
236: gb\_est167:\*  
237: gb\_est168:\*  
238: gb\_est169:\*  
239: gb\_est170:\*  
240: gb\_est171:\*  
241: gb\_est172:\*  
242: gb\_est173:\*  
243: gb\_est174:\*  
244: gb\_est175:\*  
245: gb\_est176:\*  
246: gb\_est177:\*  
247: gb\_est178:\*  
248: gb\_est179:\*  
249: gb\_est180:\*  
250: gb\_est181:\*  
251: gb\_est182:\*  
252: gb\_est183:\*  
253: gb\_est184:\*  
254: gb\_est185:\*  
255: gb\_est186:\*  
256: gb\_est187:\*  
257: gb\_est188:\*  
258: gb\_est189:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.



## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	467.4	10.6	573	139	BE755695 209602 MA
2	449.2	10.2	815	108	AU153540 AU135340
3	398.8	9.0	482	138	BE650938 UI-M-BH3-
4	371.2	8.4	954	154	BG477974 602522966
5	350.8	7.9	515	110	AV474847 AV474847
6	305.6	6.9	368	136	BE478981 163493 BA
7	263.8	6.0	117	17	A1176907 EST220513
8	250.2	5.7	466	24	A1760509 wp64d04.x
9	249	5.6	326	134	BE458694 BE458694
10	240	5.4	308	134	BB449764 BB449764
11	239.8	5.4	505	19	A1410862 A1410862
12	236.2	5.3	484	148	BF420450 UI-R-BJ2-
13	217.4	4.9	401	141	BE847128 uw22b11.Y
14	216.2	4.9	314	128	BB218602 BB218602
15	201.4	4.6	422	17	A1233991 EST230679
16	201.4	4.6	444	116	AM489638 UI-M-BH3-
17	200	4.5	400	22	A1602045 UI-R-G0-u
18	194	4.4	265	3	AA184901 m13e12.f
19	193.4	4.4	384	141	BE852200 uw22b11.x
20	193.2	4.4	247	116	AM446518 86227 MAR
21	192.4	4.4	313	128	BB219780 BB219780
22	190.2	4.3	590	166	BE368063 601221956
23	190	4.3	431	116	AM456831 UI-M-BH3-
24	174	3.9	334	163	BE097510 UI-R-BH1-
25	172.2	3.9	383	150	BF522383 UI-R-C3-S
26	158.8	3.6	927	106	AL552613 AL552613
27	158.4	3.6	3739	192	AK014835 Mus muscu
28	157	3.5	397	226	AQ270160 HS_2037.A
29	156.2	3.5	683	243	AZ426543 IM0208B09
30	152	3.4	497	15	A1072568 UI-R-C2-n
31	150.2	3.4	286	135	BB475629 BB475629
32	136.6	3.1	745	106	AL577036 AL577036
33	136.2	3.1	965	146	BF302800 602032680
34	133.2	3.0	918	106	AL556703 AL556703
35	131.8	3.0	864	18	A1324553 m192d06.Y
36	131.6	3.0	741	155	BE568567 602587557
37	131.2	3.0	942	174	BG175139 602337457
38	130.6	3.0	657	17	A1173222 u663a12.x
39	130	2.9	759	154	BG533260 601860625
40	130	2.9	1038	141	BE911604 601663080
41	129	2.9	290	141	BE848985 uw01f02.Y
42	127.8	2.9	857	175	BG242396 602354177
43	127	2.8	608	137	BE554133 ur38608.Y
44	125.4	2.8	775	175	BG246010 602358838
45	124.6	2.8			

## ALIGNMENTS

RESULT	LOCUS	DEFINITION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE
1	BE755695	209602 MARC 2BOV Bos taurus cDNA 5', mRNA sequence.	BE755695	BE755695.1	GI:10169687	cow.	1 (bases 1 to 573)	Smith,T.P.L., Casas,E., Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Laegreid 'W.M. and Keefe,J.W.	Design and use of four pooled tissue normalized cDNA libraries for

JOURNAL  
COMMENT  
EST discovery in cattle  
Unpublished (2000)  
Contact: Smith TPL  
USDA, ARS, US Meat Animal Research Center  
PO Box 166, Clay Center, NE 68933-0166, USA  
Tel: 402 762 4396  
Fax: 402 762 4390  
Email: smiththe@mail.marc.usda.gov  
Single pass sequencing. Bases called and alt\_trimmed with phred v0.980904.e. Vector identified by cross\_match with the -minscore 18 and -minmatch 12 options.  
PCR Primers  
FORWARD: AGGAACAGCATGACCAT  
BACKWARD: GTTTTCCAGTACGACG  
Plate: 59 row: N column: 5  
Seq primer: ATTTAGGTGACCTATG.  
Location/Qualifiers  
1. .573  
/organism="Bos taurus"  
/db\_xref="taxon:9913"  
/clone\_lip="MARC 2BOV"  
/tissue\_type="pooled",  
/lab\_host="DH10B"  
/note="Vector: PCMV SPORT6; Site\_1: XbaI; Site\_2: XhoI; Library made from pooled tissue from testis, thymus, semitendinosus muscle, longissimus muscle, pancreas, adrenal, and endometrium."

Query Match	10.6%	Score 467.4	DB 139	Length 573
Best Local Similarity	88.5%	Pred. No. 8.3e-123		
Matches 507	Conservative 0	Mismatches 66	Indels 0	Gaps 0
3089	ttcactgaaataatgaagaaactcagaggtccctccatccatccatccatgaagatccc	3148		
1	TTTGGCGAAGAAATGCGAAGAACTGCGGGCGCCCTCCATGTCACACAGAGACCT	60		
3149	acttttaagaagaattatcgcttccatgaacaaatgagctacataatgatatactacct	3208		
61	ACTTTTCAAGAAAGTATCGCTTCATGCAATGATGCTACGTGATGATGATGATGAT	120		
3209	ggcttagaagaatgagctcaggaatcgaaggtatcgatgatatcgctcagatggaagcaat	3268		
121	GGCTTAGTCATGGCTCACCATCAAGGATGAGTGTGCTGCTGATGCGACGACAT	180		
3269	gaaacacatccattctatcatcattcagtgacatggttcaactgtacagaaagaagagag	3328		
181	GAATATATCCATTCATTCATTTTCAGTGGCATGTGTTCACTGTGCGAAGACGAGGAG	240		
3329	tataaatggacactgtacaaatctatccaggtgtttttagagacagtggaatgttacc	3388		
241	TATTAATATGGCGGCTTACAAATCTTACCCAGGTGCTTGTGAGAGGTGGAATCTTAC	300		
3389	tcacaagctggaatttggcggttggaatgcttttttggcgagcaatcagatcggtgagat	3448		
301	TCCAAGTTGGAGCTTGGCGGATGAAATGCTTATGCGAGCACTCAAGCTGGGATG	360		
3449	agcacactttcttggttlaagaacataatgctcagactccctcgggaatgcttctga	3508		
361	AGCACTCTCTTCCTGCTGCTGATACAGCAGGAGTGTCAAAATTCACATGGGATGCTT	420		
3509	caaatagagaatttcagattacagctcaggaacataatgagacagtggtggcccaagctg	3568		
421	CGCATTTAAAGATTTTCAGATTACAGCTTCACAGACAATGTGACAGTGGGCCCAAG	480		
3569	gccagacttaatttccggatcaatcagctccggagcaccagaaggacccctttctgg	3628		
481	GCCAGACTCTATTTATTCGATCAATACGCCCTGGAGACCAAGAGATCCCTTCCTGG	540		
3629	atcaagtgatctgttgcccaatgattact	3661		



DB	541	ATCAAGGTGATCTGTTGGCCCGATGATTTAT	573
RESULT	2		
LOCUS	AUI35340	815 bp	mRNA
DEFINITION	AUI35340	PLACEL Homo sapiens cDNA clone PLACE1001786 5', mRNA sequence.	EST 24-OCT-2000
ACCESSION	AUI35340		
VERSION	AUI35340.1	GI:10995879	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 815) Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y., Yamamoto,O.,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and Isogai,T.		
TITLE	HRI human cDNA project		
JOURNAL	Unpublished (2000)		
COMMENT	Contact: Takao Isogai Genomics Laboratory Helix Research Institute 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan Tel.: 81-438-52-3951 Fax: 81-438-52-3952 Email: genomics@hri.co.jp HRI human cDNA project; 5'- & 3'-end one pass sequencing; Helix Research Institute; cDNA library construction; Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.		
FEATURES	Location/Qualifiers		
source	1..815		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="PLACE1001786"		
	/clone_1lb="PLACE1"		
	/tissue_type="Placenta"		
	/note="Vector: PME18SFL3"		
BASE COUNT	200 a 207 c 204 g 197 t	7 others	
ORIGIN			
Query Match	10.2%; Score 449.2; DB 108; Length 815;		
Best Local Similarity	98.3%; Pred. No. 1.7e-117;		
Matches	454; Conservative 0; Mismatches 8; Indels 0; Gaps 0		
QY	3788 atggcctctttggcaatgtgattcatcttgtagataaaacacatatctttaaaccctcca	3847	
Db	92 AAGGCTCTTTTGGCAATGTGATTCATCTGGATTAACACAAATATTTTAACTCCCTCCA	151	
QY	3848 attatgtcgtacatactcgtctgcaccacacattatagatcctgcgaagcactctgcg	3907	
Db	152 ATTATTGCTCGATACATCCGTTTGCACCACTCATTTATAGATTTCGACGACTCTTCCG	211	
QY	3908 atggagcttgatgggctgtgattaaatagltgcagatgcattgggaaatggagataaa	3967	
Db	212 ATGGAGTGTATGGCGCTGTATTAATTAAGTTGAGCATGTCGCAATGGGAATGGAGATAA	271	
QY	3968 gcaatatcagatgcacagattactcgtctcatcctacttaccataatgtttgcacctg	4027	
Db	272 GCAATATCGAATGCACAGATTTACTGCTTCATCTTACCAATATGTTTGCACCTCG	331	
QY	4028 tctctctcaaaagcttcgacttcaacctcccaagggaaggatgaatgcttggaagactaagg	4087	
Db	332 TCTCTCTCAAAAGCTCGACTTCACCTCCCAAGGAGGAGTAATGCCGAGACCTCAGGTG	391	
QY	4088 aataatccaaagatgtgctgcgaatggagctccagaagaacatgaaatgacagagata	4147	
Db	392 AATATTCGCAAAAGAGTGCTGCACAGTGAGCACTTCAGAAACATGAAAGTCACAGAGTA	451	
QY	4148 actactcggaggaataaatctctgcttaccagcatgtatgtgaaggagttctcctcatctcc	4207	

Db	452	ACTGACTAGGAGTAATAATCTGCTTACAGACATGATGTAAGAGACTTCCTCARTCC	511
Oy	4208	accagtcagaatgycatcagtggaactctcttttcagaat	4249
Db	512	AGCAGTCAAGATGGCCATCAGTGGACTGCTACCTTCGAAAT	553
RESULT	3		
BE650938			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
source			



described (Bonaldo, Lennon and Soares, Genome Research  
6:791-806, 1996)"

BASE COUNT 142 a 102 c 101 g 137 t

ORIGIN

Query Match 9.0%; Score 398.8; DB 138; Length 482;  
Best Local Similarity 89.2%; Pred. No. 4.2e-103;  
Matches 430; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 3740 gatggagaagatgagcactatcgaggaattccactggaacctaatggtctctt 3799  
|||||  
Db 1 GATGGGAGAAGATGCGCTGAGATTCAAGAAATCCACTGGAACTTAATGTTCTT 60  
QY 3800 ggcacatgagatctggtgataaacaatacttttaacctccaattatgctcga 3859  
|||||  
Db 61 GGCATGAGGACTGATCTGGGATTACCATPATATGTTTATCTCCAAATTAATGCTCGA 120  
QY 3860 tacatccgtttgcaacccacatcatatagatctcgacgacctctcgcatgagttatg 3919  
|||||  
Db 121 TATATCGGTTTGACACCCACTCATTTAGCATCCGTAAGTACTCTTCGATGAGATTGATG 180  
QY 3920 ggcctgatttaaatagttcagcagcatgaggaattggaatggaatgaagaatacagat 3979  
|||||  
Db 181 GCGTGATTTAAACAGCTTCAGCATACCATTTGGAAATGGAAGTAATATATCATGAT 240  
QY 3980 gcacagattactgcttactactactactaacaatagtcttcacacctgcttccctcaaa 4039  
|||||  
Db 241 ACACAAATCACTGCTCTATCTACTTCAACAACTGTTCTTACTTGTCTCTTCACAA 300  
QY 4040 gctcgacttcaacctcaagaagagagatgctcgagagacctgagtgaaataccaaa 4099  
|||||  
Db 301 GCTGACTTCACTCCCGGGAAGAGACTAATGCCGCGGACTCAGTGATGATCCAAA 360  
QY 4100 gagtgcgtcgaagtgcctccagaagaacaatgaagtcacaggaataactactcagga 4159  
|||||  
Db 361 CAATGCTTGAAGAGTCACTTCAAAAAGCAATGAATCTACTGGAATTAATCCCAAGGA 420  
QY 4160 gtaaatctctgcttaccagcatgatagtgaagagttctctcatctccacagtcagat 4219  
|||||  
Db 421 GTGAAATCTCTTTACACAGATGTTGTGAAGAGTCTTATTTTACGACAGTCAAGAT 480  
QY 4220 gg 4221  
||  
Db 481 GG 482

RESULT 4  
BG477974 954 bp mRNA EST 21-MAR-2001  
LOCUS 602522966F1 NIH\_MGC\_20 Homo sapiens cDNA clone IMAGE:4641352 5',  
DEFINITION mRNA sequence.  
ACCESSION BG477974.1 GI:13410253  
VERSION BG477974.1  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 954)  
AUTHORS NIH-MGC http://mgs.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: ATCC/DCPD/DPF  
CDNA Library Preparation: Ling Hong/Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLCM1408 row: c column: 17

FEATURES High quality sequence stop: 626.  
Location/Qualifiers  
1..954

source

/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4641352"  
/clone\_lib="NIH-MGC\_20"  
/tissue\_type="melanotic melanoma"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: skin; Vector: pOTB; Site: 1: XhoI; Site: 2:  
EcoRI; cDNA made by oligo-dr priming. Directionally  
cloned into EcoRI/XhoI sites using the following 5'  
adapter: GGCACGAG(c). Size selected >500bp for average  
insert size 1.8kb. Library constructed by Ling Hong in  
the laboratory of Gerald M. Rubin (University of  
California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 233 a 276 c 232 g 212 t

ORIGIN

Query Match 8.4%; Score 371.2; DB 154; Length 954;  
Best Local Similarity 86.5%; Pred. No. 4.8e-95;  
Matches 585; Conservative 0; Mismatches 68; Indels 23; Gaps 15;

QY 105 atatgcaagtgatctcggtgagctgctgtgagcgaagaattccctcagaagtccaa 164  
|||||  
Db 108 AGATGACCGAGATGCCGCGGAGAAACGTCGCCGGAATTCCTCTAGAGTGCAA 167  
QY 165 aactttccatccaacacctcagtcgtgacaaaagaactcgtttgtagaattcacg 224  
|||||  
Db 168 AATCTTTTCATTCACACCTCAGTGTGTACAAAAGACTGTTGTGATTCACGG 227  
QY 225 ttcaaccttcaaatcgtcgaagcgaagccacctggaatggatctgtagtctcaca 284  
|||||  
Db 228 ATCACTTTTCAACATCGCTTAACGCAAGCGCACCTGAGTGGCTGCTAGTCTACCA 287  
QY 285 ttcaggtgaggtgtttatgatacagtgctacttaactaagaacatgagcttccatctg 344  
|||||  
Db 288 TCCAGGCTGAGGTTTATGATACAGTGTGTACCTTAAGAACATGCGTCCCATCTTG 347  
QY 345 tcaagcttcatctgtgtgtgatacctactcgaagaagctctga--"ggagactgaataga 402  
|||||  
Db 348 TCAGTCTCATGCTGTGTGTGTATCCAGTGAAGCTTGTGACGGAGCCTGAATATGA 407  
QY 403 tgaatcagaccagtcaca--"ggagagaagaagatgata-aagttctccgtgtggaagca 459  
|||||  
Db 408 TGATCAGACCGACGTCACAAACGGGAGACAGAAAGATGATACAGTCTTCCCTGGAAGCCA 467  
QY 460 tacat-atgtctgagaggtctg---aaagaatggtccaatgagctctgaccactgt 515  
|||||  
Db 468 TACCTCATGCTGTGGCAGGCTCTCAAGAGACATGTCACCTGCGCTCTGACCCACTGA 527  
QY 516 gcc--ttacctactcatatcttctcatgtggaactg--taaaagacttgaat--tcagc 571  
|||||  
Db 528 TGCCATTACTACTCATATATCTTCTCATGTGAGACCTGCGTCAAAAGCTTAATCTCAGGA 587  
QY 572 ctcaattggagcctactagatgta--gagaaggagtcgccaaggaana--gacaag 628  
|||||  
Db 588 CTCATTGAGACCTCTACTAGATATGATAGAGCAGGAGTCTGGCCAGCAACCGACACAG 647  
QY 629 accttttcaaa-aatttatctactt--tttggtatttgatggaaggaana--gttgga 684  
|||||  
Db 648 ACCTTGACACCATATATCTTCTGTGGCTGTATGTATGTAACGGAAACGTTGGCA 707  
QY 685 ctcaagaana--caagaactcctgatgcag--gataaggaatgctgcatctgctggagcctg 741  
|||||  
Db 708 CTAGAAAGAGAGAACTCTGATGAGAGCGATAGAGATGCTGCATGCTGCTGGGGCCT 767  
QY 742 gctcaaatgacaca 757  
|||  
Db 768 GGCCTTACATTTGCACA 783



RESULT	5
LOCUS	AV748467
DEFINITION	AV748467 NPC Homo sapiens cDNA clone NPACXA05 5', mRNA sequence.
ACCESSION	AV748467
VERSION	AV748467.1 GI:10906315
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE	1 (bases 1 to 515)
AUTHORS	Song,H., Peng,Y., Gu,Y., Yang,Y., Gao,G., Xiao,H., Xu,X., Li,N., Qian,B., Liu,F., Qu,J., Gao,X., Cheng,Z., Xu,Z., Zeng,L., Xu,S., G. W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Ye,M., Zhang,Q., Han,Z., Chen,Z., Hu,R. and Chen,J.
TITLE	Homo sapiens NPC library cDNA clones
JOURNAL	Unpublished (2000)
COMMENT	Contact: Qinghua Zhang Shanghai Institute of Endocrinology, Rui-Jin Hospital 197 Rui-jin II Road, Shanghai 200025, P. R. China Tel.: 86-21-64370045(ex.663332) Fax: 86-21-64743206 Email: mshlims.stn.sh.cn This clone is available at Shanghai Hematology Institute in Shanghai. Chinese National Human Genome Center at Shanghai 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong.

FEATURES	SOURCE	Location/Qualifiers
		1. .515
		/organism="Homo sapiens"
		/db_xref="taxon:9606"
		/clone="NPCAXA05"
		/clone_lib="NPC"
		/tissue_type="pituitary"
		/dev_stage="Adult"
		/lab_host="SOLR"
		/note="Vector: pbluescript sk(-); Site_1: EcoRI; Site_2: XhoI"
BASE COUNT		132 a 114 c 131 g 135 t 3 others
ORIGIN		

Query Match	7.9%	Score	350.8	DB	110	Length	515
Best Local Similarity	95.3%	Pred.	No. 2.7e-89				
Matches	383	Conservative	0	Mismatches	14	Indels	5
						Gaps	2

QY	3788	atgctctctcttggcaatgtgagttcactcgggaataaaacaatatctttaaaccctca	3847
Db	111	AAGGTCTCTTCTTGGCAATGTGGATCTACTCGGGATTAACACAAATATTTTAAACCTTCA	170
QY	3848	atattgctcgatcatalcaccgtcttgcaaccaactcatalagcatctgcagcactctgc	3907
Db	171	ATTATTCGTGATATCATCCGTTTGGCAOCCACATATTATAGCATTTTGGCAGCACTTTGGC	230
QY	3908	atggagatgbatbggctgtgtgattaaatagtttcagcatalgcatttggaaatgagataa	3967
Db	231	ATGAGATGTGATGGCCTGTGATTTAAATAGTTCAGCATGCCATTGGGATGAGAGATAAA	290
QY	3968	gcaatalcagatcagacagatatactgcttcatctctaacttaccaatatgtttgcaccctg	4027
Db	291	GCAATATTCAGATGACAGATTACTGCTTCACTCTTAAACAAATATGTTTCCACCTGG	350
QY	4028	tctctcttcaaaagctgaacttcacctcctcaaggagagatatactgcctgagacctcaagtg	4087
Db	351	TCCTCTTCAAAAGCTGCACCTTCAOCTTCAAGGAGGAGTAATCTCTGGAGACTTAAGTG	410
QY	4088	ataataccaaaagtgctgcgcaatgagacttcacagaagacatatgaatcaacagagata	4147
Db	411	AATATATCCAAAGAGTGGCTTCAAGTGCATTTTCAGAAAGCAATGTGTCACA-NGAAGTA	469
QY	4148	actactcagagtgataaaatctctgcttctaccagcatgatgtg	4189

Db 470 ACTACTTAGGAGTA---CTCTTCTTACCAGCATGTATGTG 507

RESULT	6						
BE478981		BE478981	368 bp	mRNA	EST	28-AUG-2000	
LOCUS		1636433	BARC	5BOV Bos taurus	CDNA 5'	mRNA sequence.	
DEFINITION		BE478981					
ACCESSION		BE478981.1	GI:9598514				
VERSION		EST.					
KEYWORDS							
SOURCE		COW.					

ORGANISM  
Bos laurus  
Elkayotze; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea  
Bovidae; Bovinae; Bos.  
1 (bases 1 to 368)

**AUTHORS** Sonstegard, T.S., Capucio, A.V., Van Tassell, C.P., Astwell, M.S. and Wells, K.D.

**TITLE** Mapping of Expressed Sequence Tags from a normalized bovine mammary gland cDNA library

**JOURNAL** unpublished (2000)

**COMMENT** Contact: Sonstegard TS  
USDA, ARS, Beltsville Agricultural Research Center  
Bldg. 200 Rm 2A, Beltsville, MD 20705, USA  
Tel: 301 504 8416  
Tel: 301 504 8416  
Fax: 301 504 8414  
Email: tads@psi.barc.usda.gov  
Single pass sequencing. Bases called and alt trimmed with phred  
v0.960904.c. Vector identified by cross\_match with the -mismatch 18  
and -mismatch 12 options.

FORWARD: AGGAACACGTTATGACCAT  
 BACKWARD: GTTTCCTCAGTACAGAG  
 Plate: 141 row: 6 column: 5  
 Seq primer: ATTAGTGACACTRATAG.

FEATURES	Location/Qualifiers
source	1..368
	/organism="Bos taurus"
	/db_xref="taxon:9913"
	/clone_lib="BARC 5B0V"
	/tissue_type="pooled"
	/lab_host="DH10B"
	/note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI; library made from pooled mRNA isolated from mammary tissues at eight physiological, developmental, and disease states."
BASE COUNT	90 a 87 c 98 g 93 t
ORIGIN	

[illegible]



Db 241 GCCAGACTTATTTCTGTGATCAATCACCCCTGAGCACCAGATCCCTCTCTGG 300

Qy 3629 atcaagtgatctgttggcccaatgattatcaagcatcaagaccaggtgcccgt 3688  
|||||  
Db 301 ATCAAGTGATCTGTGGCCCGCATGATTAATCAACATCTGACTCAGGGGCCGG 360

Qy 3689 cagaagtt 3696  
|||||

Db 361 CAGAAGTT 368

RESULT 7  
A1176907/c 517 bp mRNA EST 20-JAN-1999  
LOCUS EST220513 Normalized rat ovary, Bento Soares Rattus sp. cDNA clone  
DEFINITION ROBY42.3' end, mRNA sequence.  
ACCESSION A1176907  
VERSION A1176907.1 GI:3727545  
KEYWORDS EST.  
SOURCE Rattus sp.  
ORGANISM Rattus sp.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
1 (bases 1 to 517)  
Lee,N.H., Glodek,A., Chandra,I., Mason,T.M., Quackenbush,J.,  
Kerlavage,A.R. and Adams,M.D.  
Rat Genome Project: Generation of a Rat EST (RESt) Catalog & Rat  
Gene Index  
Unpublished (1998)  
JOURNAL Contact: Lee, NH  
COMMENT The Institute for Genomic Research  
9712, Medical Center Drive, Rockville, MD 20850, USA  
Tel: (301)-838-3529  
Fax: (301)-838-0208  
Email: nhlee@ligr.org  
Seq primer: M13-21.  
FEATURES  
source  
1..517  
/organism="Rattus sp."  
/db\_xref="ATCC (Inhost):2031594"  
/db\_xref="taxon:10118"  
/clone="ROBY42"  
/clone\_lib="Normalized rat ovary, Bento Soares"  
/note="Organ: ovary; Vector: pT733pac; Site\_1: EcoRI;  
Site\_2: NotI"

BASE COUNT 93 a 174 c 154 g 96 t

ORIGIN

Query Match 6.0%; Score 263.8; DB 17; Length 517;  
Best Local Similarity 70.9%; Pred. No. 2.9e-64;  
Matches 365; Conservative 0; Mismatches 147; Indels 3; Gaps 1;

Qy 2474 cagaagcccgagcttcaaaaacacagacacatttattctgtcagtgagag 2533  
|||||  
Db 513 CAGGAGACCCGGCGCTTCAGACAAAACGGGACATTTTCATCGCGCCCGTAAGAG 454

Qy 2534 ctctggattatggatgagtagctcccaatgttcaagaacagaggtcagagtgagc 2593  
|||||  
Db 453 CTTTGGACATACAGGGCGTGTCTCGCTGAGACTCCGGAGACAGGGCGTGAAGTGGG 394

Qy 2594 agtgcctcaagttcaagaagttgtttccaggaatttctgtagtgcttcttaccag 2653  
|||||  
Db 393 GACCCAGCTAGGTTCGGAAAGTGTTCGTTCAAGGACGGGCTGTTACGACGAG 334

Qy 2654 cccctatacgttgagagaaatgaatgaacatttgagactcctgggacatatagaagca 2713  
|||||  
Db 333 CGGCTGACCGGTGGGAGCTGAGACGACACCTGGGGCTACTGGGCCATATACGGGGC 274

Qy 2714 gaagtgaagataatataatcgttaacttcaagaatacagagcctctgctctatccttc 2773  
|||||  
Db 273 GAAGTGAAGACAAACATCGTGTGACCTTCGGAAACAGGCGCTCCGCCCTACTCTTC 214

Qy 2774 tattcagccctatttcttattaggaatcagaagcgaaggagcagaacctagaaaaaac 2833  
|||||  
Db 213 TACTTCACCTCTGTGTCTTACCAGGACACGCT---GGAGTGCACCAAGAGCAAC 157

Qy 2834 ttgtcaagcctaaatgaacaaacttacttcttgaaagtgaacatcatatgtgaccc 2893  
|||||  
Db 156 TTGTGGCGCCCAACGAGACCAAAACTTACTTTCGAGAGTCGGGCCACATGCGGCC 97

Qy 2894 actaaagtatgattgactgcaagcctgggtactattctctgattgttgacctgaaaaa 2953  
|||||  
Db 96 ACGGACGCGGAGTTTGACTGCAAGCGCTGATTTCTCCAGCTGAGACTCGAGCA 37

Qy 2954 gattgacatcagggcctgattgagccctctgt 2988  
|||||  
Db 36 GACCTGCACCTGGGCTGATTTGGCCACTGCTCGT 2

RESULT 8  
A1760509/c 466 bp mRNA EST 20-DEC-1999  
LOCUS wh64004.x1 NCI-CGAP\_K1d11 Homo sapiens cDNA clone IMAGE:2385511 3'  
DEFINITION similar to gb:M14113 COAGULATION FACTOR VIII PRECURSOR (HUMAN);,  
mRNA sequence.  
ACCESSION A1760509  
VERSION A1760509.1 GI:5176176  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 466)  
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncigap.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
JOURNAL Contact: Robert Strausberg, Ph.D.  
COMMENT Email: cgaps-remail.nih.gov  
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.  
Emmert-Buck, M.D., Ph.D.  
cDNA Library Preparation: M. Bento Soares, Ph.D.  
cDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNLT at:  
www-dio.llnl.gov/dbp/image/image.html  
Insert Length: 1564 Std Error: 0.00  
Seq primer: -40UP from Gibco  
High quality sequence stop: 404.  
FEATURES  
source  
1..466  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2385511"  
/clone\_lib="NCI-CGAP\_K1d11"  
/lab\_host="DH10B"  
/note="Organ: kidney; Vector: pT733-Pac (Pharmacia) with  
a modified polylinker. Site\_1: Not I; Site\_2: Eco RI;  
Plasmid DNA from the normalized library NCI-CGAP\_K1d1 was  
prepared, and ss circles were made in vitro. Following HAP  
purification, this DNA was used as tracer in a subtractive  
hybridization reaction. The driver was PCR-amplified cDNAs  
from a pool of 5,000 clones made from the same library  
(clones 1322376-1323911, 1456007-1456775, and  
1500552-1502855). Subtraction by Bento Soares and M.  
Fatima Bonaldo."

BASE COUNT 128 a 96 c 77 g 165 t

ORIGIN

Query Match 5.7%; Score 250.2; DB 24; Length 466;  
Best Local Similarity 95.2%; Pred. No. 2.3e-60;  
Matches 258; Conservative 0; Mismatches 13; Indels 0; Gaps 0;



QY	1197	aaacttgggtacattacattgcttgcctgaagagagagactggagactatgcttccttaagcttc	1256
Db	453	AAACTTTGGGTACATTACATTGCTGCTGCAAGAGAGAGAGACTGGAGCTATGCTCCCTTAATGCC	394
QY	1257	tcgcccccgatgcagagaagttlaataaagtcacatatttgacaacatgycgcccacgagatg	1316
Db	393	TGCGCCCGGATGACAGAAAGTTATTAAGTCATATTGGAACAATGGGCCCTCAGCGGATGG	334
QY	1317	gtagaagaatcaaaaaaagtcgcgattatggcctacacagatgaacaccttaagactcgtg	1376
Db	333	GTAGAAGATCAAAAAAGTCGATTTATGCGCTACACACAGATGAACACTTTAAGACTCGCG	274
QY	1377	aagcctatcagatgaatcagaagactctggagcccttaatttgatgggaagcttggagaca	1436
Db	273	AAGCTATTTCAGCATTAATCAGAAATCTTGGGACCTTTACTTTATGAGGAGATTGGAGACA	214
QY	1437	cactgttgattatatttaagaatcaagcaag	1467
Db	213	CACCTGTGGTAGTTGAAGAAAAAGATTTTAG	183
RESULT	9		
LOCUS	BB458694		
DEFINITION	BB458694 RIKEN full-length enriched, 12 days embryo spinal ganglion	EST	21-JUL-2000
VERSION	BB458694	Mus musculus cDNA clone D13008B20.3, similar to I05573 Mus	
KEYWORDS	domesticus coagulation factor VIII mRNA, mRNA sequence.		
ACCESSION	BB458694		
VERSION	BB458694.1	GI:9354187	
KEYWORDS	EST.		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. I (bases 1 to 326)		
AUTHORS	Kono, H., Aikawa, K., Akahira, S., Akiyama, J., Arahawa, T., Carninci, P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N., Hirozane, T., Hori, F., Ishii, Y., Ishikawa, J., Ishikawa, T., Itoh, M., Izawa, M., Kadota, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N., Kiyosawa, H., Kojima, Y., Kondo, S., Koya, S., Kurihara, C., Kusakabe, M., Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y., Ono, T., Owa, C., Saito, H., Sakai, C., Sato, K., Shibata, K., Shibata, Y., Shigemoto, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Sugahara, Y., Suzuki, H., Suzuki, H., Tagawa, A., Takahashi, F., Tomimaga, N., Toya, T., Tsunoda, Y., Watabiki, A., Watanabe, S., Yamamura, T., Yamana, I., Yano, R., Yasunishi, A., Yokota, T., Yoshida, K., Yoshiki, A., Yoshino, M., Muramatsu, M., and Hayashizaki, Y.		
TITLE	RIKEN Mouse ESTs (Kono, H., et al.)		
JOURNAL	Unpublished (2000)		
COMMENT	Contact: Yoshinori Hayashizaki Genome Exploration Research Group, Life Science Tsukuba Center, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan Tel: +81-298-36-9013 Fax: +81-298-36-9098 Email: genome-res@rc.riken.go.jp URL: http://genome.riken.go.jp/ Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoka, S., Sasakikawa, N., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y. Thermosensitization and thermoinactivation of thermolabile enzymes by trichostatin and its application for the synthesis of full length cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998) Itoh, M., Kikuchi, N., Akiyama, J., Shibata, K., Izawa, M., Kawai, J., Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki, Carninci, P., and Hayashizaki, Y. Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999) High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)		

Please visit our web site (<http://genome.rtc.riken.go.jp>) for further details.

Location/Qualifiers  
1. 326

FEATURES  
source

```

BASE COUNT
ORIGIN
      87 a      82 c      56 g      101 t
      plbuescript KS(+) after bulk excision from lambda FLC I."
      /organism="Mus musculus"
      /db_xref="taxon:10090"
      /clone="D130058E20"
      /clone_1lb="RIKEN full-length enriched, 12 days embryo
      spinal ganglion"
      /tissue_type="spinal ganglion"
      /dev_stage="12 days embryo"
      /lab_host="DH10B"
      /note="Site_1: SalI; Site_2: BamHI; cDNA library was
      prepared and sequenced in Mouse Genome Encyclopedia
      Project of Genome Exploration Research Group in Riken
      Genomic Sciences Center and Genome Science Laboratory in
      RIKEN. Division of Experimental Animal Research in Riken
      contributed to prepare mouse tissues. 1st strand cDNA was
      primed with a primer [5',
      GAGGAGAGAGGCGGCCGACACGAGTTTGTGTGTGTGTGTVN 3'], cDNA was
      prepared by using trehalose thermo-activated reverse
      transcriptase and subsequently enriched for full-length by
      cap-trapper. Second strand cDNA was prepared with the
      primer adapter of sequence [5',
      GAGGAGAGAGATTCTCGAGTTCAATTAATTAATATACCCCCCCCCC 3']. cDNA
      was cleaved with BamHI and XhoI. Vector: a modified
      pBluescript KS(+) after bulk excision from lambda FLC I."

```

Query Match	5.6%	Score 249	DB 134	Length 326
Best Local Similarity	86.0%	Pred. No. 4.3e-60		
Matches 276	Conservative 0	Mismatches 45	Indels 0	Gaps 0

BASE COUNT	87 a	82 c	56 g	101 t
ORIGIN				

Oy 740 tggcctaataatgcacacagtcgaattgttatgtaaacaggctctctccaggcttgatcyga 799  
|| ||| ||||||| |||| | | ||||| ||| |||||||  
Db 2 TGCCCTTTTATCCACACTTTCAAAGGCATCTTAACAGCCTTTTCCAGGCGTGATTGGA 61

Oy 800 tgcacagaaatcagtcattatggcatgtgatggaatggcaccaactctctaagtgcac 859

Db 62 TACCATAGGAATCAGTCTACTACGGCACGTGATTGGAATGGGCTCCACTCCTCGAATACAC 121

Qy 860 tcaatatccctcgaaagtcacacattctctgtgaggaaccatgcgcagcgctcttgaa 919  
|||||  
Db 122 TCATATTTCTCGAAGGTACACATTTTTTTGTAGGGAACCAACCCCTCAAGCTTCATTGCGAG 1811

QY	920 atctcgcaataactttccttactgtccaacacctctgatgacctggagagtctcta 979
Db	182 ATATCACCACAATACTTTCCTTAAGTCGTTCAACACTCTTGATAGATCTTGGCGAGTTCCCTA 2411

DQ 980 ctgcttctgcataatccttcaccacaacatgatgcatgtgaagcttatgtccaagttagac 1033

DB 242 cttatttttttcatttctctcccaataaacatgatggcatggaagctttatgtccaagtgat 3011

QY	1040	agctgtccagagaaacccaa	1060
DB	302	AGCTGCCCTGAGGAATCCCAA	322

RESULT 10  
BB449764

LOCUS	508 bp	mtDNA	E31	21 JUL 2000
DEFINITION	BB449764	RIKEN full-length enriched, 9 days embryo	Mus musculus	
	CDNA clone D030073H24	3' similar to U05573	Mus domesticus	
	COAGULATION FACTOR VIII mRNA, mRNA SEQUENCE.			

ACCESSION	BB449704
VERSION	BB449764.1
KEYWORDS	GI:9310799
SOURCE	house mouse.
EST.	

ORGANISM  
Mus musculus  
Eukaryota; Metazoa;  
Mammalia; Eutheria;  
Rodentia; Sclurognathi;  
Muridae; Mus.



1 (bases 1 to 308)  
Konno, H., Aizawa, K., Akahira, S., Akiyama, J., Arakawa, T., Carninci

The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center

Location/Qualifiers  
1. .308

Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoka, S., Sasaki, N., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y. Thermostabilization and thermoactivation of thermostable enzymes by trehalose and its application for the synthesis of full length cDNA. *Proc. Natl. Acad. Sci. U.S.A.* 95 (2), 520-524 (1998)

Itoh, M., Kitsuana, T., Akiyama, J., Shibata, K., Iwaw, M., Kawai, J., Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki, Y., and Hayashizaki, Y.

Automated filtration-based high-throughput plasmid preparation system. *Genome Res.* 9 (5), 463-470 (1999)

Carninci, P. and Hayashizaki, Y.

High-efficiency full-length cDNA cloning. *Methods Enzymol.* 303, 19-44 (1999)

Please visit our web site (<http://genome.rtc.riken.go.jp>) for further details.

85 a	80 c	50 g	93 t
------	------	------	------

Query Match	5.4%;	Score 240;	DB 134;	Length 308;
Best Local Similarity	86.2%;	Pred. No. 1.6e-57;		
Matches 261; Conservative	0;	Mismatches 35;	Indels 0;	Gaps 0;
QY	765	gtttgtaaaagctccttcgcgaagtcctgatggatgccacagaataatcaagtattggc	824	
I				

Db	9	GCATGTAAACAGTCTCTCTCCAGTTCTGTATTGGATCCCATGGAATACAGTCTACTGGC	68
Qy	825	atgtgatattggaatggygcacaccccttcgaagtgcactcaatatalctctgaaagttcaacat	884
Db	69	ACCGATGTGGAATGGGACACCTCCGGAATACATCAATATTTCTCGAAGGTCCACAT	128
Qy	885	ttctctgtaggaacaacatcgcgcagcgcttcctcttgaaatcgcgcaataacttctctactg	944
Db	129	TTTTTTTGAGAGAACCAOCCCTCAAGCTTCATGTGAGATATCACCAATTAATTCTTCTACTG	188
Qy	945	ctcaaacactcttgatgagaccttgaaagttctactglttgtcatatalctctccacc	1004
Db	189	CTCAAAACACTCTTGATGATCTTGGCCAGTCTCTACTATTTTTCATTCTCTCTCCATA	248
Qy	1005	aacacttgatgcatggaagactatgtccaagatgaacagctgctccagagaacccca	1060
Db	249	AACATGATGCGATGGAAGACTTATTTCAAGTGAATATGCTCCCTTGAGGAATGCCAA	304

Unpublished (1998)  
Contact: Lee, NH

REFERENCE  
1 (bases 1 to 505)  
Lee,N.H., Glodek,A., Chandra,I., Mason,T.M., Quackenbush,J.,  
Kerlavage,A.R. and Adams,M.D.  
Rat Genome Project: Generation of a Rat EST (RESE) Catalog & Rat  
Gene Index

Location/Qualifiers  
1. .505

```

/organism="Rattus sp."
/db_xref="taxon:10118"
/clone="RHEC09"
/clone_1p="Normalized rat heart, Bento Soares"
/note="Organ: heart; Vector: pT73pac; site_1: EcORI
Site_2: NotI"
BASE COUNT      83 a      148 c      165 g      109 t
ORIGIN

```

Similarity 69.6%; Pred. No. 2.3e-57; 0; Mismatches 147; Indels 9; Gaps 2

<b>OY</b>	736	gacctggcctaataatgcacacagfcaatlgttatgttaaaccggtctctcgccagtctgat	791
<b>Dd</b>	505	GGCCCAAGGCCACGSGTGCACACTGTGAATGGCTACACGACC GGACACTCGCGGGTGTTAC	444
<b>OY</b>	796	tgaattgccacaaggaaatcagtcataatggcatlgtatgynaatyggcacactcctbaagt	851
<b>Dd</b>	445	CGGATGGCTTCGGAACATCCTCGAAGTAAGTAGCATGATGCGTGGGACACACGCCAGACAT	381
<b>OY</b>	856	gacataaatatctcttgtaagtgaaatatcttctgttagaacatctgcgcggctctt	911
	385	ACACTCATCTCTCTCGAAGGCCACACATTCTCTGTGTGGAGCCACCGTCAAGGCTTCTCT	320















GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 18, 2001, 04:30:38 ; Search time 11937.4 Seconds  
(without alignments)  
3773.245 Million cell updates/sec

Title: US-09-689-430-1\_COPY\_150\_4914

Perfect score: 4765

Sequence: 1 cctcttcaagtaaacagta.....ttgggtcgttgcgcgac 4765

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 10228115 segs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST.\*

1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_est3:\*  
4: gb\_est4:\*  
5: gb\_est5:\*  
6: gb\_est6:\*  
7: gb\_est7:\*  
8: gb\_est8:\*  
9: gb\_est9:\*  
10: gb\_est10:\*  
11: gb\_est11:\*  
12: gb\_est12:\*  
13: gb\_est13:\*  
14: gb\_est14:\*  
15: gb\_est15:\*  
16: gb\_est16:\*  
17: gb\_est17:\*  
18: gb\_est18:\*  
19: gb\_est19:\*  
20: gb\_est20:\*  
21: gb\_est21:\*  
22: gb\_est22:\*  
23: gb\_est23:\*  
24: gb\_est24:\*  
25: gb\_est25:\*  
26: gb\_est26:\*  
27: gb\_est27:\*  
28: gb\_est28:\*  
29: gb\_est29:\*  
30: gb\_est30:\*  
31: gb\_est31:\*  
32: gb\_est32:\*  
33: gb\_est33:\*  
34: gb\_est34:\*  
35: gb\_est35:\*  
36: gb\_est36:\*  
37: gb\_est37:\*  
38: gb\_est38:\*  
39: gb\_est39:\*  
40: gb\_est40:\*  
41: gb\_est41:\*  
42: gb\_est42:\*  
43: gb\_est43:\*

44: em\_esthum10:\*  
45: em\_esthum11:\*  
46: em\_esthum12:\*  
47: em\_esthum13:\*  
48: em\_esthum14:\*  
49: em\_esthum15:\*  
50: em\_esthum16:\*  
51: em\_esthum17:\*  
52: em\_esthum18:\*  
53: em\_esthum19:\*  
54: em\_esthum20:\*  
55: em\_esthum21:\*  
56: em\_esthum22:\*  
57: em\_esthum23:\*  
58: em\_esthum24:\*  
59: em\_esthum25:\*  
60: em\_esthum26:\*  
61: em\_esthum27:\*  
62: em\_esthum28:\*  
63: em\_esthum29:\*  
64: em\_esthum30:\*  
65: em\_esthum31:\*  
66: em\_esthum32:\*  
67: em\_esthum33:\*  
68: em\_esthum34:\*  
69: em\_esthum35:\*  
70: em\_esthum36:\*  
71: em\_esthum37:\*  
72: em\_esthum38:\*  
73: em\_esthum39:\*  
74: em\_esthum40:\*  
75: em\_esthum41:\*  
76: em\_esthum42:\*  
77: em\_esthum43:\*  
78: em\_esthum44:\*  
79: em\_esthum45:\*  
80: em\_esthum46:\*  
81: em\_esthum47:\*  
82: em\_esthum48:\*  
83: em\_esthum49:\*  
84: em\_esthum50:\*  
85: em\_esthum51:\*  
86: em\_esthum52:\*  
87: em\_esthum53:\*  
88: em\_esthum54:\*  
89: em\_esthum55:\*  
90: em\_esthum56:\*  
91: em\_esthum57:\*  
92: em\_esthum58:\*  
93: em\_esthum59:\*  
94: em\_esthum60:\*  
95: em\_esthum61:\*  
96: em\_esthum62:\*  
97: em\_esthum63:\*  
98: em\_esthum64:\*  
99: em\_esthum65:\*  
100: em\_esthum66:\*  
101: em\_esthum67:\*  
102: em\_esthum68:\*  
103: em\_esthum69:\*  
104: em\_esthum70:\*  
105: em\_esthum71:\*  
106: em\_esthum72:\*  
107: em\_esthum73:\*  
108: em\_esthum74:\*  
109: em\_esthum75:\*  
110: em\_esthum76:\*  
111: em\_esthum77:\*  
112: em\_esthum78:\*  
113: em\_esthum79:\*  
114: em\_esthum80:\*  
115: em\_esthum81:\*  
116: em\_esthum82:\*



117: gb\_est48:\*  
118: gb\_est49:\*  
119: gb\_est50:\*  
120: gb\_est51:\*  
121: gb\_est52:\*  
122: gb\_est53:\*  
123: gb\_est54:\*  
124: gb\_est55:\*  
125: gb\_est56:\*  
126: gb\_est57:\*  
127: gb\_est58:\*  
128: gb\_est59:\*  
129: gb\_est60:\*  
130: gb\_est61:\*  
131: gb\_est62:\*  
132: gb\_est63:\*  
133: gb\_est64:\*  
134: gb\_est65:\*  
135: gb\_est66:\*  
136: gb\_est67:\*  
137: gb\_est68:\*  
138: gb\_est69:\*  
139: gb\_est70:\*  
140: gb\_est71:\*  
141: gb\_est72:\*  
142: gb\_est73:\*  
143: gb\_est74:\*  
144: gb\_est75:\*  
145: gb\_est76:\*  
146: gb\_est77:\*  
147: gb\_est78:\*  
148: gb\_est79:\*  
149: gb\_est80:\*  
150: gb\_est81:\*  
151: gb\_est82:\*  
152: gb\_est83:\*  
153: gb\_est84:\*  
154: gb\_est85:\*  
155: gb\_est86:\*  
156: gb\_est87:\*  
157: gb\_est88:\*  
158: gb\_est89:\*  
159: gb\_est90:\*  
160: gb\_est91:\*  
161: gb\_est92:\*  
162: gb\_est93:\*  
163: gb\_est94:\*  
164: gb\_est95:\*  
165: gb\_est96:\*  
166: gb\_est97:\*  
167: gb\_est98:\*  
168: gb\_est99:\*  
169: gb\_est100:\*  
170: gb\_est101:\*  
171: gb\_est102:\*  
172: gb\_est103:\*  
173: gb\_est104:\*  
174: gb\_est105:\*  
175: gb\_est106:\*  
176: gb\_est107:\*  
177: gb\_est108:\*  
178: gb\_est109:\*  
179: gb\_est110:\*  
180: gb\_est111:\*  
181: gb\_est112:\*  
182: gb\_est113:\*  
183: gb\_est114:\*  
184: gb\_est115:\*  
185: gb\_est116:\*  
186: gb\_est117:\*  
187: gb\_est118:\*  
188: gb\_est119:\*  
189: gb\_est120:\*

190: gb\_est121:\*  
191: gb\_est122:\*  
192: gb\_est123:\*  
193: gb\_est124:\*  
194: gb\_est125:\*  
195: gb\_est126:\*  
196: gb\_est127:\*  
197: gb\_est128:\*  
198: gb\_est129:\*  
199: gb\_est130:\*  
200: gb\_est131:\*  
201: gb\_est132:\*  
202: gb\_est133:\*  
203: gb\_est134:\*  
204: gb\_est135:\*  
205: gb\_est136:\*  
206: gb\_est137:\*  
207: gb\_est138:\*  
208: gb\_est139:\*  
209: gb\_est140:\*  
210: gb\_est141:\*  
211: gb\_est142:\*  
212: gb\_est143:\*  
213: gb\_est144:\*  
214: gb\_est145:\*  
215: gb\_est146:\*  
216: gb\_est147:\*  
217: gb\_est148:\*  
218: gb\_est149:\*  
219: gb\_est150:\*  
220: gb\_est151:\*  
221: gb\_est152:\*  
222: gb\_est153:\*  
223: gb\_est154:\*  
224: gb\_est155:\*  
225: gb\_est156:\*  
226: gb\_est157:\*  
227: gb\_est158:\*  
228: gb\_est159:\*  
229: gb\_est160:\*  
230: gb\_est161:\*  
231: gb\_est162:\*  
232: gb\_est163:\*  
233: gb\_est164:\*  
234: gb\_est165:\*  
235: gb\_est166:\*  
236: gb\_est167:\*  
237: gb\_est168:\*  
238: gb\_est169:\*  
239: gb\_est170:\*  
240: gb\_est171:\*  
241: gb\_est172:\*  
242: gb\_est173:\*  
243: gb\_est174:\*  
244: gb\_est175:\*  
245: gb\_est176:\*  
246: gb\_est177:\*  
247: gb\_est178:\*  
248: gb\_est179:\*  
249: gb\_est180:\*  
250: gb\_est181:\*  
251: gb\_est182:\*  
252: gb\_est183:\*  
253: gb\_est184:\*  
254: gb\_est185:\*  
255: gb\_est186:\*  
256: gb\_est187:\*  
257: gb\_est188:\*  
258: gb\_est189:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.



## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	467.4	9.8	573	139	BE755695	209602 MA
2	449.2	9.4	573	108	BE755695	209602 MA
3	398.8	8.4	482	138	BE650938	UT-M-BH3-
4	371.2	7.8	954	134	BE477974	602522966
5	350.8	7.4	515	110	AV748467	AV748467
6	305.6	6.4	368	136	BE478981	163493 BA
7	263.8	5.5	517	17	BE716907	EST220513
8	250.2	5.3	466	24	BE760509	wh64404 x
9	249	5.2	326	134	BE458694	BB458694
10	240	5.0	308	134	BE449764	BB449764
11	239.8	5.0	505	19	BE410862	EST239155
12	236.2	5.0	484	148	BE420450	UT-R-B32-
13	217.4	4.6	401	141	BE847128	uw22b11.y
14	216.2	4.5	314	128	BE218602	BR218602
15	201.4	4.2	422	17	BE233991	EST230679
16	201.4	4.2	444	116	BE489638	UT-M-BH3-
17	200	4.2	400	22	BE602045	UT-R-GO-u
18	194	4.1	265	3	BE184901	mc93e12.f
19	193.4	4.1	384	141	BE852200	uw22b11.x
20	193.2	4.1	247	116	BE446518	BE6227 MAR
21	192.4	4.0	313	128	BE219780	BE219780
22	190.2	4.0	590	166	BE368063	601221956
23	190	4.0	431	116	BE456831	UT-M-BH3-
24	174	3.7	334	163	BE097510	UT-R-B01-
25	172.2	3.6	383	150	BE522383	UI-R-C3-S
26	158.8	3.3	927	106	BE552613	AL552613
27	158.4	3.3	3739	192	BE4014835	Mus muscu
28	157	3.3	397	226	BE0270160	HS-2037_A
29	156.2	3.3	683	243	BE426543	IM0208B09
30	152	3.2	497	15	BE1072568	UI-R-C2-n
31	150.2	3.2	286	135	BE475629	BE475629
32	136.6	2.9	745	106	BE577036	AL577036
33	136.2	2.9	965	146	BE302800	602032680
34	133.2	2.8	918	106	BE556703	AL556703
35	131.6	2.8	864	18	BE1324533	mt92606.y
36	131.6	2.8	807	108	BE134073	BE134073
37	131.2	2.8	741	155	BE658567	602857557
38	130.6	2.7	942	174	BE175139	BE175139
39	130	2.7	657	17	BE1173222	wd63812.x
40	130	2.7	759	154	BE653260	601860625
41	129	2.7	1038	141	BE911604	601663080
42	127.8	2.7	290	141	BE848985	uw01f02.y
43	127.2	2.7	775	175	BE246010	602358838
44	127	2.7	857	175	BE242396	602354177
45	125.4	2.6	608	137	BE554133	ur38e08.y

## ALIGNMENTS

RESULT	1	LOCUS	BE755695	573 bp	mrna	EST	15-SEP-2000
DEFINITION	209602 MARC 2B0V Bos taurus CDNA 5', mRNA sequence.						
ACCESSION	BE755695						
VERSION	BE755695.1						
KEYWORDS	EST.						
SOURCE	COW.						
ORGANISM	Bos taurus						
REFERENCE	1 (Bases 1 to 573)						
AUTHORS	Smith,T.P.L., Casas,E., Stone,R.T., Heaton,M.P., Grose,W.M., Bennett,G.A., Fahrmeir,S.C., Freking,B.A., Rohrer,G.A., Laegreid ,W.W. and Keele,J.W.						
TITLE	Design and use of four pooled tissue normalized cDNA libraries for						

## JOURNAL COMMENT

EST discovery in cattle

Unpublished (2000)

Contact: Smith TPL

USDA, ARS, US Meat Animal Research Center

PO Box 166, Clay Center, NE 68933-0166, USA

Tel: 402 762 4366

Fax: 402 762 4390

Email: smith@email.marc.usda.gov

Single pass sequencing. Bases called and alt-trimmed with phred

v0.980904.e. Vector identified by cross-match with the -minscore 18

and -mismatch 12 options.

PCR Primers

FORWARD: AGGAACAGCTATGACCAT

REVERSE: GTTTCACGATGACGACG

Plate: 59 row: N column: 5

Seq primer: ATTAGTGACACTAAG.

Location/Qualifiers

1. 573

/organism="Bos taurus"

/db\_xref="taxon:9913"

/clone\_lib="MARC 2B0V"

/library="pooled"

/lab\_host="DH10B"

/note="Vector: pCMV SPORT6; Site\_1: XbaI; Site\_2: XhoI;

Library made from pooled tissue from testis, thymus,

adrenal, and endometrium."

BASE COUNT 149 a 134 c 146 g 144 t

## ORIGIN

## BASE COUNT

## ORIGIN

## BASE COUNT

## ORIGIN

## BASE COUNT

## ORIGIN

## BASE COUNT

## ORIGIN

## BASE COUNT

## ORIGIN

## BASE COUNT

## ORIGIN

## BASE COUNT

## ORIGIN

## BASE COUNT

## ORIGIN

## BASE COUNT

## ORIGIN

## BASE COUNT

## ORIGIN

## BASE COUNT

## ORIGIN

## BASE COUNT

## ORIGIN

## BASE COUNT

## ORIGIN

## BASE COUNT

## ORIGIN

## BASE COUNT

## ORIGIN

## BASE COUNT

## ORIGIN

## BASE COUNT

## ORIGIN

## BASE COUNT



DB	541	ATCAAGCTGATCTGTGGCGCCGATGATTAT	573
RESULT	2		
LOCUS	AUI35340	815 bp	mRNA
DEFINITION	AUI35340	PLACE1 Homo sapiens CDNA clone	EST 24-OCT-2000
ACCESSION	AUI35340	PLACE1001786	5', mRNA
VERSION	AUI35340.1	GI:10955879	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Okuyama, T., Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 815)		
AUTHORS	Ota, T., Nishikawa, T., Suzuki, Y., Ishii, S., Saito, K., Kawai, Y., Yamamoto, Y., Wakamatsu, A., Nakamura, Y., Nagai, T., Sugano, S. and Isogai, T.		
TITLE	HRI human CDNA project		
JOURNAL	Unpublished (2000)		
COMMENT	Contact: Takao Isogai		
	Genomics Laboratory		
	Helix Research Institute		
	1532-3 Yana, Kisarazu, Chiba 292-0812, Japan		
	Tel: 81-438-52-3851		
	Fax: 81-438-52-3952		
	Email: genomics@hri.co.jp		
	HRI human CDNA project; 5'- & 3'-end one pass sequencing; Helix Research Institute; CDNA library construction; Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute		
FEATURES	Location/Qualifiers		
Source	1..815		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="PLACE1001786"		
	/clone_lib="PLACE1"		
	/tissue_type="placenta"		
	/note="Vector: pME18SFL3"		
BASE COUNT	200 a 207 c 204 g 197 t		
ORIGIN	7 others		

Db	452	ACTGCTAGGAGAGTAAATCTCTCTCTTACGACATGATATGTGTAAGAGATTCTCTCATCTCC	511
QY	4477	agcagctcaagatggccatcagctggactctcttttcagaat	4518
Db	512	AGCAGTCAAGATGGCCATCATGATGGACTGCTACCTTCGATT	553
RESULT	3		
LOCUS	BE650938		
DEFINITION	BE650938	482 bp	EST
ACCESSION	U1-M-BH3-asn-g-04-0-0-I .1	NIH_BMAP_M_S4	Mus musculus cDNA clone
VERSION	U1-M-BH3-asn-g-04-0-0-I 5'		mRNA sequence.
KEYWORDS	BE650938		
SOURCE	BE650938.1	GI:976762	EST.
ORGANISM	house mouse.		
REFERENCE	Mus musculus		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
TITLE	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
JOURNAL	1 (bases 1 to 482)		
MEDLINE	Bonaldo,M.F., Lennon,G. and Soares,M.B.		
COMMENT	Normalization and subtraction: two approaches to facilitate gene discovery		
	Genome Res. 6 (9), 791-806 (1996)		
	9704477		
	Contact: Chin, H		
	National Institute of Mental Health		
	6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD		
	20892-9643, USA		
	Tel: 301 443 1706		
	Fax: 301 443 9890		
	Email: mestr@mail.nih.gov		
	cDNA Library Preparation: M.B. Soares Lab Clone distribution:		
	Researchers may obtain BMAP cDNA clones from RESEARCH GENETICS. It		
	should be noted that Bento Soares is generating a small number of		
	additional specialized non-redundant arrays of BMAP cDNAs whose		
	availability will be compared under appropriate and limited		
	collaborative arrangements		
	Seq primer: M13 Reverse.		
FEATURES	Location/Qualifiers		
source	1..482		

	Query Match	9.4%	Score 449.2	DB 108	Length 815
	Best Local Similarity	98.3%	Pred. M. 5.4e-119		
	Matches 454	Conservative	0	Mismatches 8	Indels 0
				Gaps	0
QY	4057	atgctctcttggcgaatgtagatcatccttggataataaacacaaatctttaaaccctcca	4116		
Db	92	AAGGCTCTTCTTTGGCAATGATGATTCATCTGGATTAACACAAATTTTAAACCTCCA	151		
QY	4117	attatgtcgtatcatccgtcttgcaccaactaatatagaatccgaagcactctgcg	4176		
Db	152	ATTATGTCTCGATACATCGTTTGGACCCACCACTAATATAGATTGCGACGACTCTTCGC	211		
QY	4177	atggatgtgaatggagctgtgatattaatagtticaagatgcccatttggaaatgaagttaa	4236		
Db	212	ATGGAGTGTATGGCTGTGATTTAAATATCTTCACAGCATCCATTTGGAAATGGAGATTAA	271		
QY	4237	gcaaatcagaatgcacagaattactcttcatacttaccatattglttgcacctga	4296		
Db	272	GCAATATCAGATGCACAGATTACTGCTTATCTTACTTACCAATATGTTTGCACCTGG	331		
QY	4297	tctcttcaaaaagctcgaacttcaacctccaagagagagtaatgactctgaagactcaagt	4356		
Db	332	TCTCTTCAAAAAGCTCGACTTCACCTCCCAAGGAGAGTAATGGCTCGAGACTCAGGTG	391		
QY	4357	aataatccaaaagatgctcgaagtgaatccagaagaacaatgaagtccagaagta	4416		
Db	392	AATATATCCAAAAGAGTGGCTGCAAGTGAGCTTCCAGAAACATGATGAATCCACAGAGTA	451		
QY	4417	actactcaggaggtataaatctctgcttaccagcagtatagtgaaggagttcctcatctcc	4476		

/db\_xref="taxon:10090"  
 /clone="U1-M-BH3-asm\_g-04-0-U1"  
 /clone\_lib="NIH\_BMAP\_M\_S4"  
 /dev\_stage="27-32 days"  
 /lab\_host="DH10B (Life Technologies)"  
 /note="Vector: pTR73D-pac (Pharmacia) with a modified  
 polylinker: Site\_1: Not I; Site\_2: Eco RI; The  
 NIH\_BMAP\_M\_S4 library is a substracted library of a series  
 ultimately derived from a mixture of individually tagged  
 normalized libraries from ten regions of the mouse brain  
 (cerebellum, brain stems, olfactory bulbs, hypothalamus,  
 cortex, amygdala, basal ganglia, pineal gland, striatum,  
 hippocampus) after a series of subtractions to reduce the  
 representation of cDNAs from which ESTs had already been  
 generated. The following serially substracted libraries  
 were generated in this process: NIH\_BMAP\_M\_S4,  
 NIH\_BMAP\_M\_S3.3, NIH\_BMAP\_M\_S3.2, NIH\_BMAP\_M\_S3.1,  
 NIH\_BMAP\_M\_S2, NIH\_BMAP\_M\_S1. The substracted library  
 (NIH\_BMAP\_M\_S4) was constructed as follows: PERamplified  
 cDNA inserts from NIH\_BMAP\_M\_S3.3, NIH\_BMAP\_M\_S3.2, and  
 NIH\_BMAP\_M\_S3.1 clones from which 3' ESTs had been derive  
 was used as a driver in a hybridization with a pool of  
 the NIH\_BMAP\_M\_S3.3, NIH\_BMAP\_M\_S3.2, and NIH\_BMAP\_M\_S3.1  
 libraries in the form of single-stranded circles. The  
 remaining single-stranded circles (substracted library)  
 was purified by hydroxypatite column chromatography,  
 converted to double-stranded circles and electroporated  
 into DH10B bacteria (Life Technologies) to generate the  
 NIH\_BMAP\_M\_S4 library. This procedure has been previousl



described (Bonaldi, Lennon and Soares, Genome Research 6:791-806, 1996)  
 BASE COUNT 142 a 102 c 101 g 137 t  
 ORIGIN

Query Match 8.4%; Score 398.8; DB 138; Length 482;  
 Best Local Similarity 89.2%; Pred. No. 2e-104;  
 Matches 430; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 4009 gatgggaagaatgagcactatcgaggaattccactgaaaccttaagtctctt 4068  
 |||||  
 Db 1 GATGGGAAGAATGAGCTGAGTATCAAGAAATTCACCTGAACTTATGCTTTT 60  
 QY 4069 ggcaatgagattatctgggataaacaatttttaacctccattatgctga 4128  
 |||||  
 Db 61 GGCATGAGACTCATCTGCGATTAAGCATATAGTTTAAATCCCAATATGCTCGA 120  
 QY 4129 tacatcgtttgcaaccacattatagcattcgagcactcttcgcatgagttgag 4188  
 |||||  
 Db 121 TATATCGCTTGGACACCCACATCTCTAGCATCCGTACTCTTCGCGATGAGTTGATG 180  
 QY 4189 ggcgtgtatctaataagttgcagcattgaggaatgagagaaagacatacagat 4248  
 |||||  
 Db 181 GCGTGTGATTTAAACAGTTGACATACCATGGGAATGAAAGTAAATATCAAGAT 240  
 QY 4249 gcaagattactgcttaccacttaccattatgctgacactgcttccttcaaa 4308  
 |||||  
 Db 241 ACACAAATCACTGCTCATCTCTACTTACCAACATGTTTGTCTTCTCTTCAACA 300  
 QY 4309 gctcgactcaccctccaaaggaggaatgctgagagcactcagtgtaataatc 4368  
 |||||  
 Db 301 GCTGACTTACCTCCACGAGGAGGACATATGCTGCGACCTCAGCTGAATGATCCAAA 360  
 QY 4369 gattggtgcgaatgagcttccagaagaacaaatgaagtcacagagtaactc 4428  
 |||||  
 Db 361 CAATGGTTCAGAGTGAATTTACAAAGAACATGAAAGTCACTGGAATTAACCCAGGA 420  
 QY 4429 gtaaatctcgttaccagattatgtaaggagttcccatctccagagtaagat 4488  
 |||||  
 Db 421 GTGAATCTCTTTACAGCATGTTGTGAAGAAGTCTTATTTTCAAGCATAGAT 480  
 QY 4489 gg 4490  
 ||  
 Db 481 GG 482

RESULT 4  
 BG477974 954 bp mRNA EST 21-MAR-2001  
 LOCUS 60252966F1 NIH\_MGC\_20 Homo sapiens cDNA clone IMAGE:4641352 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BG477974  
 VERSION BG477974.1 GI:13410253  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 REFERENCE 1 (bases 1 to 954)  
 NIH-MGC http://mgi.nci.nih.gov/.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: ggapds@email.nih.gov  
 Tissue Procurement: ATCC/DCTD/DRP  
 CDNA Library Preparation: Ling Hong/Rubin Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.llnl.gov  
 Plate: LNCM1408 row: c column: 17

FEATURES  
 source  
 High quality sequence, stop: 626.  
 Location/Qualifiers  
 1..954

/organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4641352"  
 /clone\_id="NIH\_MGC\_20"  
 /tissue\_type="melanotic melanoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: skin; Vector: pOTB7; Site: 1: XhoI; Site: 2:  
 EcoRI; cDNA made by oligo-dT priming. Directionally  
 cloned into EcoRI/XhoI sites using the following 5'  
 adaptor: GGCAAGAG(G). Size-selected >500bp for average  
 insert size 1.8kb. Library constructed by Ling Hong in  
 the laboratory of Gerald M. Rubin (University of  
 California, Berkeley) using ZAP-cDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies)."  
 BASE COUNT 233 a 276 c 232 g 212 t  
 ORIGIN

Query Match 7.8%; Score 371.2; DB 154; Length 954;  
 Best Local Similarity 86.5%; Pred. No. 3e-96;  
 Matches 585; Conservative 0; Mismatches 68; Indels 23; Gaps 15;

QY 374 atatgcaaatgagctgagctgagctgagcgaagattccctcctgaagtgcga 433  
 |||||  
 Db 108 AGATGCACCGAGATGCGGCTGCGCAAAAGCGTCCGGGATTTCTCTAGAGTCCAA 167  
 QY 434 aatctttccattcaaccctcagctgctgtacaaaagactctgtttgataattcag 493  
 |||||  
 Db 168 AATCTTTTCCATTCAACACTCAGTCGCTGTAACAAAAGACTCTGTTGATGATCACGG 227  
 QY 494 ttcaactttcaaatcgtctgaagccaagccctgtagtggctgctgaagttcctaca 553  
 |||||  
 Db 228 ATCACTTTTCAACATCTCCTTAAGCCCAAGCCACCTGGATGGGTCTGATGCTTACCA 287  
 QY 554 tccagagctgaggttatatatacagatggttcattacattaaagaatgcttccactg 613  
 |||||  
 Db 288 TCCAGGCTGAGGTTTATATACAGTGTGATTAACCTTAAGAACATGGCTTCCATCTCG 347  
 QY 614 tcaagcttcacatgctgtgtatcctactcgtgaagctcttga--ggagctgaatatga 671  
 |||||  
 Db 348 TCAGTCTTACATGCTGTGTGTATCTGTAAGCTTCTGACGGAGACGTGATATGA 407  
 QY 672 tgatcagaccagtcaca--ggagagaagaagatgata--aagcttccctggtgagaagcca 728  
 |||||  
 Db 408 TGATCAGACCAAGTCMAAGGGGAGAAAGAGATGATATCAAGCTTCCCTGGTGGAAAGCCA 467  
 QY 729 taat-atgtctgagagtcctg--aaagagaatggtccaaatggccttgaaccactgt 784  
 |||||  
 Db 468 TACCTCATGTCTGGAGGTCTGCAAAAGAGATGTCACCTGGCTCTGACCCACAGGA 527  
 QY 785 gcc--taccactacatacttctcattcattgtagacctg--taaaagactgaat--tcaagc 840  
 |||||  
 Db 528 TGCCATTACCTACATCATATCTTTCTCATGTGAGACCTGCTCAAGACTTGAATCACA 587  
 QY 841 ctcaattgagccctactagatgta--gagaaggagatgctggccaagaaa--gacacag 897  
 |||||  
 Db 588 CTCATTGGAGGCCCTACATGTATGATGACGAGGAGTGTGGCGAGGCAAAACCGACACAG 647  
 QY 898 accttgacaca--aatataactact--tttgctgtatgtagaagagaaa--gttggca 953  
 |||||  
 Db 648 ACCTTGACACACCATCTATACTACTTGTGGCTGTATATGATGAACGGGAACGTTGGCCA 707  
 QY 954 ctcaagaa--caagaactccttgatgag--gataggagatgctgcatcgtcggacctg 1010  
 |||||  
 Db 708 CTCAGAAACAGCAAGAACTCTTGTATGAGGAGATGAGATGCTGATGCTCGGGGCT 767  
 QY 1011 gctcaaatgacaca 1026  
 |||||  
 Db 768 GGCCTTACATGTCACA 783



RESULT	5
AV748467	
LOCUS	AV748467 515 bp mRNA EST 19-OCT-2000
DEFINITION	AV748467 NPC Homo sapiens cDNA clone NPCCXA05 5', mRNA sequence.
ACCESSION	AV748467
VERSION	AV748467.1 GI:10906315
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 515) Song,H., Peng,Y., Gu,Y., Yang,Y., Gao,G., Xiao,H., Xu,X., Li,N., Qian,B., Liu,F., Qu,J., Gao,X., Cheng,Z., Xu,Z., Zeng,L., Xu,S., G., W., Tu,Y., Jia,T., Fu,G., Ren,S., Zhong,M., Lu,G., Ye,M., Zhang,Q, , Han,Z., Chen,Z., Hu,R. and Chen,J. Homo sapiens NPC library CDNA clones Unpublished (2000) Contact: Qinghua Zhang Shanghai Institute of Endocrinology, Rui-Jin Hospital 197 Rui-Jin II Road, Shanghai 200025, P. R. China Tel.: 86-21-64370045(ex. 663532) Fax: 86-21-64743206 Email: mbshl@ems.stn.sh.cn This clone is available at Shanghai Hematology Institute in Shanghai. Chinese National Human Genome Center at Shanghai 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong. Location/Qualifiers 1..515 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="NPCAXA05" /clone_lib="NPC" /tissue_type="pituitary" /dev_stage="Adult" /lab_host="SOLR" /note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2: XhoI"
TITLE	
JOURNAL	
COMMENT	
FEATURES	
SOURCE	
BASE COUNT	132 a 114 c 131 g 135 t 3 others
ORIGIN	

	Query Match	7.4%	Score 350.8	DB 110	Length 515
	Best Local Similarity	95.3%	Pred. No. 2e-90		
	Matches 383	Conservative	0	Mismatches 14	Indels 5
				Gaps 2	
QY	4057	atgcgtctcttggcgaatgtagatcatctgtagataaacaacataatttcaacctcca	4116		
Db	111	AAGGCTCTTGTGGCAATGCTGGATTCATCTGGGATTAACACAAATATTTTAACCTTCA	170		
QY	4117	attattgtcgalatataccgthttgacccaactaatatagcatctgcagcaactcttcgc	4176		
Db	171	ATTATGTCTGCATACATCCGTTTGACCCAACTCATTTATAGCATTCGCAGCACTCTTGC	230		
QY	4177	atgagatgtgaaggcgtcgtgattaaatagtgtgcagcagtcgcaatttggaaatggaatgtaa	4236		
Db	231	ATGGAGTGTGATGGCGCTGTGATTTAAATAGTTGCCACATATGCCATTTGGGATGGAGGTAA	290		
QY	4237	gcaatcatagatgcagcagattactgcttcatactcctaacttaccataatgthttgcaccttg	4296		
Db	291	GCAATATATAGATGCACAGATTACTGCTTCATCTCTTATCCAAATATGTTTCCCACTGG	350		
QY	4297	tctctctcaaaagctcgacttcaacctccaagggaggagtaatgcttgcgtagagactcaagtg	4356		
Db	351	TCTCTCTCAAAAGCTGCAGCTTACCTTCAAGGAGGAGATTAATGCTCGAGACTCTCAGTG	410		
QY	4357	aataatccaaaagatgtgctgcagtgcaagtgtaactccagaagacaaatgaagaatccaggaatga	4416		
Db	411	AATATATCCAAAAGAGATGGCTTTCATAGTGGCACTTTCAGAAAGACAAATNGTGCAC-NGAAGTA	469		
QY	4417	actactcaggaagtataaactctgctgttaccagcatgatatgtg	4458		

Db	470	ACTACTTAAGAGATA----	CTCTCTTACAGCATGATG	507
RESULT	6			
LOCUS	BE478981	368 bp	mRNA	EST
DEFINITION	163493 BARC 5BOV Bos taurus CDNA 5', mRNA sequence.			28-AUG-2000
ACCESSION	BE478981			
VERSION	BE478981.1	GI:9598514		
KEYWORDS	EST.			
SOURCE	COW.			
ORGANISM	Bos taurus			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea; Bovidae; Bovinae; Bos.			
AUTHORS	1 (bases 1 to 368)			
TITLE	1 (bases 1 to 368)			
COMMENT	Mapping of Expressed Sequence Tags from a normalized bovine mammary gland cDNA library			
JOURNAL	Unpublished (2000)			
CONTACT	Contact: Sonstegard TS			
USDA, ARS, Beltsville Agricultural Research Center				
Bldg. 200 Rm 24, Beltsville, MD 20705, USA				
Tel: 301 504 8416				
Fax: 301 504 8414				
Email: tads@psi.barc.usda.gov				
Single pass sequencing. Bases called and alt_trimmed with phred v0.980904.e. Vector identified by cross_match with the -mnscore 18 and -mismatch 12 options.				
PCR Primers				
FORWARD: AGGAACAGCTATGACCAT				
BACKWARD: GTTTCGCAGTCAGCAG				
Plate: 141 row: G column: 5				
Seq primer: ATTAGGTGACATCATAG.				
Location/Qualifiers				
1..368				
/organism="Bos taurus"				
/db_xref="taxon:9913"				
/clone_lib="BARC 5BOV"				
/tissue_type="pooled"				
/lab_host="DH10B"				
/note="Vector: pCMV SPORT6; Site.1: XbaI; Site.2: XhoI; library made from pooled mRNA isolated from mammary tissues at eight physiological, developmental, and disease states."				
BASE COUNT	90 a	87 c	98 g	93 t
ORIGIN				
Query Match	6.4%; Score 305.6; DB 136; Length 368;			
Best Local Similarity	89.4%; Pred. No. 2.5e-77;			
Matches 329; Conservative 0; Mismatches 39; Indels 0; Gaps 0;				
3598	tataaatgacacgtgtacaatctctatccagcgtgtttttsagacagtggaagtgtacc	3657		
1	TGTAAATGGCGGCTCTACATCTACCCAGGTGCTTTTGAGACGTTGGAATGCTACCA	60		
3658	tccaaagtgtgaattttgscgggtggaatgcttatgtgcagacatctacatgctggatg	3717		
61	TTCCAAAGTTGGGACTTGGCGGATGGAATGCTTATTTGGCGAGCACCTTACAGCTGGGATG	120		
3718	agcacacttttcctcgtgttacagacaatagtgtcagactcccccctggagatgctctgga	3777		
121	AGCACTCTCTTCCTGCGTGTACAGCAGGAGAGTCAATTCACATGGGATGCTTCTGGA	180		
3778	cacattagagatttcagattacagcttcagagacaatagtgcagtggtggcccaagctg	3837		
181	CGCATTTAGAGATTTTTCAGATTACAGCTTACAGACAAATTTGACAGTGGGCCCCCAACTG	240		
3838	ggcagactcatcttcgcgtacataatcctcggagaccacaaggagcccttctcttg	3897		



LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT	FEATURES	BASE COUNT	ORIGIN
241	GCCACACTTCATTATTCGATGATCAATCAACGGCTGGAGACCAAGATGCCCTCTCTTGG	300												
QY	3898	atcaagtgatgacgtgttgaccacaatgatattcaacgcatcaagaaccagggtgccgt	3957											
Db	301	ATCAGATGGAGATCTGTGGCGCGCATGATATTATTCACAGCATCTCTGACTCAGGGTGCCCG	360											
QY	3958	cagaagtc 3965												
Db	361	CAGAAGTT 368												
RESULT	7													
LOCUS	A1176907/c	517 bp	mRNA	EST	20-JAN-1999									
DEFINITION	EST220513	Normalized rat ovary,	Bento Soares	Rattus sp.	cDNA clone									
ACCESSION	ROVBY42.3	end, mRNA sequence.												
VERSION	A1176907													
KEYWORDS	EST.	GI:3727545												
SOURCE	Rattus sp.													
ORGANISM	Rattus sp.													
REFERENCE	1 (bases 1 to 517)													
AUTHORS	Lee,N.H., Glodok,A.,	Chandra,I.,	Mason,T.M.,	Quackenbush,J.,										
TITLE	Rat Genome Project: Generation of a Rat EST (RESE) Catalog & Rat													
JOURNAL	Unpublished (1998)													
COMMENT	Contact: Lee, NH													
	The Institute for Genomic Research													
	9712, Medical Center Drive, Rockville, MD 20850, USA													
	Tel: (301)-838-3529													
	Fax: (301)-838-0208													
	Email: nhlee@tigr.org													
	Seq primer: M13-21.													
FEATURES	Location/Qualifiers													
source	1..517													
	/organism="Rattus sp."													
	/db_xref="ATCC (inhost):2031594"													
	/db_xref="taxon:10118"													
	/clone="ROVBY42"													
	/clone_1db="Normalized rat ovary, Bento Soares"													
	/note="Organ: ovary; Vector: pUT33Pac; site_1: EcoRI;													
	site_2: NotI"													
BASE COUNT	93 a	174 c	154 g	96 t										
ORIGIN														
Query Match	5.5%	Score 263.8;	DB 17;	Length 517;										
Best Local Similarity	70.9%;	Pred. No. 4.2e-65;												
Matches 365;	Conservative 0;	Mismatches 147;	Indels 3;	Gaps 1;										
QY	2743	cagagcccccgcagcttccaaagaacaacagacactatttattgctgcagatggaag	2802											
Db	513	CAGGACCCCGCGGCTTTCAGACGAAACCGCGCACTATTTCATGCGCGCCCTGAAGAG	454											

Oy	3043	tattctagccttatttctctaaggaagaatcaggagcaagacgaaactagaaaaaac	3102
Db	213	TACTCCACCTCGTGTCTCCTACCCTGGAAACCGACGCT---	66AGCTGCACCAAGAGCAC
Oy	3103	tttgcaagcctaataatgcacaacaaacttacttlttgaaagtgcacaacatcatatgtcaccc	3162
Db	156	TTTGTCGCCGCCAACGAGACCAAAACTTATTCTGTGAGAGTCCGGCCCCACATGGCGCCC	97
Oy	3163	actaaagatgatgttgactgcgaagcctggcgcttatctctatgltgtacctgtgaaaaa	3222
Db	96	ACGGACGGCGAGTTTGACTGCAAGGCCCTGTATTTCTCCGAGCTGACCTCGAGCGA	37
Oy	3223	gatgtcacctcaggcgcctatttggaacccttgtgt	3257
Db	36	GACCTGCACCTCGGGCTCATTTGGCGCCACTGCTCGT	2
RESULT	8		
LOCUS	A1760509/c		
DEFINITION	A1760509	466 bp	mRNA
DESCRIPTION	wh64d04.xl NCI_CGAP_Kid11 Homo sapiens CDNA clone IMAGE:2385511 3'		
KEYWORDS	similar to gp:M4113 COAGULATION FACTOR VIII PRECURSOR (HUMAN); , mRNA sequence.		
VERSION	A1760509		
SOURCE	A1760509.1 GI:5176176		
ORGANISM	human.		
REFERENCE	Homo sapiens		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap. National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index		
JOURNAL	Unpublished (1997)		
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nh.gov Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emert-Buck, M.D., Ph.D. CDNA Library Preparation: M. Bento Soares, Ph.D. CDNA Library Arrayed by: Greg Lennon, Ph.D. DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html Insert Length: 1564 Std Error: 0.00 Seq primer: -40UP from Gibco High quality sequence stop: 404. Location/Qualifiers		
FEATURES	source		
	1..466	/organism="Homo sapiens"	,
	/db_xref="taxon:9606"		
	/clone_image="2385511"		
	/clone_id="NCI_CGAP_Kid11"		
	/lab_host="DH10B"		
	/note="Organ: kidney; Vector: pTZ19-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; Plasmid DNA from the normalized library NCI_CGAP_Kid3 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (clones 1322376-1323911, 1456007-1456775, and 1500552-1502855). Subtraction by Bento Soares and M. Fatima Bonaldo."		
BASE COUNT	128 a	96 c	77 g 165 t
ORIGIN			
Query Match	5.3%	Score 250.2;	DB 24; Length 466;
Best Local Similarity	95.2%;	Pred. No. 3;7e-61;	
Misest 258; Conservative	0;	Mismatches 13;	Indels 0; Gaps 0;



QY	1466	aaacttggtacattacatctgctgctgaagaaagagcagcgagcctatgctccttaagcc	1525
DB	453	AAACTTGGGTACATTACATTGCTGCTGAAGAAGAGACTGGGACTATGCTCCCTTAAGTCC	394
QY	1526	tcgcccccatgtacacgaagtctataaagtcacatcttgacacaatgagccctcagcgatg	1585
DB	393	TGCGCCCGCATGTACAGAAAGTTATPATAAGTCATPATTGTGAACAATAGGCCCTCAGCGAATTG	334
QY	1586	gtagggaagtacaaaaaagtcgcgatttatggcattacacagatgaacaccttaagactcgt	1645
DB	333	GTAGGAAGATCAAAAAAGCCGATTTATGCGATTACACACATGATAAACCTTTAAGACTCGG	274
QY	1646	aagcattcagatgaatcaagaactccttgagaccttaacttatggagggaagtggagaca	1705
DB	273	AAGCATATTCAGCATGAATCAGCAATCTTGCGACCTTTACTTTATGGGCAAGTTGGAGACA	214
QY	1706	cactgttgattatatttaagaatcagcaag	1736
DB	213	CACGTGTGGTAGTTGAAGAAAAAGATTTTAAAG	183
RESULT	9		
LOCUS	BB458694		
DEFINITION	BB458694 RIKEN full-length enriched, 12 days embryo spinal ganglion Mus musculus cDNA clone D13008E20.3 similar to U05573 Mus domesticus coagulation factor VIII mRNA, mRNA sequence.		
VERSION	BB458694.1	GI:9354187	
KEYWORDS	EST.		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 326)		
AUTHORS	Kono, H., Aikawa, K., Akahira, S., Akiyama, J., Arakawa, T., Carninci, P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N., Hirozane, T., Hori, F., Ishii, Y., Ishikawa, J., Ishikawa, T., Itoh, M., Izawa, M., Kadota, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N., Kiyosawa, H., Kojima, Y., Kondo, S., Koya, S., Kurihara, C., Kusabe, M., Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, T., Ono, T., Owa, C., Saito, H., Sakai, C., Sato, K., Shibata, K., Shibata, Y., Shigemoto, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Sugihara, Y., Suzuki, H., Suzuki, H., Tagawa, A., Takahashi, F., Tomimaga, N., Toya, T., Tsunoda, Y., Watabiki, A., Watanabe, S., Yamamura, T., Yamanaka, I., Yano, R., Yasunishi, A., Yokota, T., Yoshida, K., Yoshiki, A., Yoshino, M., Muramatsu, M., and Hayashizaki, Y.		
TITLE	RIKEN Mouse ESTs (Kono, H., et al.)		
JOURNAL	Unpublished (2000)		
COMMENT	Genome Exploration Research Group, Life Science Tsukuba Center, Genome Science Laboratory The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center 3-1-1 Koyada, Tsukuba, Ibaraki 305-0074, Japan Tel: +81-298-36-9013 Fax: +81-298-36-9098 Email: genome-resetc.riken.go.jp/ URL: http://genome.etc.riken.go.jp/ Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoka, S., Sasakawa, N., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y. Thermostabilization and thermocatalytic activation of the modifiable enzymes by triethanolamine and its application for the synthesis of full length cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998) Itoh, M., Kikunishi, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J., Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki, Y., and Hayashizaki, Y. Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999) Carninci, P. and Hayashizaki, Y. High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)		

FEATURES					
Source					
Location/Qualifiers					
1..326					
/organism="Mus musculus"					
/db_xref="taxon:10090"					
/clone="D130058E20"					
/clone_lib="RIKEN full-length enriched, 12 days embryo					
/spinal_ganglion"					
/tissue_type="spinal ganglion"					
/dev_stage="12 days embryo"					
/lab_host="DH10B"					
/note="Site.1: Salt; Site.2: BamHI: cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5']					
GAGGAGAGGAGCGCCGCACTGCAGTGGTTTGTATTTTTTTTNN 3'}. cDNA was prepared by using trihalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5']					
GAGGAGAGATTTCGACGTTAAATTAATTAATCCTCCCCCCC 3'}. cDNA was cleaved with BamHI and XhoI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLX I."					
BASE COUNT      87 a            82 c            56 g            101 t					
ORIGIN					
Query Match	5.2%	Score 249;	DB 134;	Length 326;	
Best Local Similarity	86.0%;	Pred. No. 7.1e-61;			
Matches 276;	Conservative	0;	Mismatches 45;	Indels 0;	Gaps 0;
Oy	1009	tggcctaaatgcacacagtcagtgtatgtaaacaggctcctgccagtgctgatgga	1068		
Db	2	tgccttttatccacacttttcaaaaggccattctaagacctctttccagtgctgatgga	61		
Oy	1069	tgcacaggaatcatcgctctatgcatgcatgtagaatgagaccactctgaatgac	1128		
Db	62	taccataagaatcacgtctactgacgcacgngattggaatggcgtccacnctgaatatcac	121		
Oy	1129	tcaatatccctcgaagtgcacacattctcttgagagaaccatcgccagcgctcttgaa	1188		
Db	122	tcaaatattccctggaaggtccacacatttttgtgagagAACACCCTCAAGCTTCATTGGAG	181		
Oy	1189	atctgcacaataactctctactgctcctcaaacacccctgtagtacctggagacgttta	1248		
Db	182	atattCACCATTAACCTTCTTACTGCTCAACAACCTTGATGATCTTGGGAGTTCCTA	241		
Oy	1249	ctgttttgcatalatctctcccacaacaatgcatgcatgtagaactlatgccaagttagc	1308		
Db	242	ctatttttttcattttctctccatTAACAATGAATGAGCATGANGAAGCTTAGTCAAAATGAT	301		
Oy	1309	agctgtccaagagaacccccaa	1329		
Db	302	AgCTGCCCTCGAGGAATCCCA	322		
RESULT_10					
LOCUS					
BB449764					
DEFINITION					
BB449764 RIKEN full-length enriched, 9 days embryo Mus musculus					
CDNA clone DO30073H24.3 similar to L05573 Mus domesticus					
cosimulation factor VIII mRNA, mRNA sequence.					
ACCESSION					
BB449764					
VERSION					
BB449764.1 GI:9310799					
KEYWORDS					
EST.					
house mouse.					
SOURCE					
Mus musculus					
ORGANISM					
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.					







Accession	Source	Organism	Version	Keywords	Definition	LOCUS	Result
1185	ggaatctcgcgaataactcttccttaacgctgcacaaacactcttgatggaccttgaaagct	1244					
335	GGAGATCTCACCATTAGACTCTGCTCACCAGCGCTCAGACGCTGGCTGATGGACTCGGGCGGTT	266					
1245	tctactgtttgttcatactctctctccacaacaatgatgacatggaagcttatgtcaaaagt	1304					
265	CCTGCTCTTCTGTGCACACTCTCCTCCACACAGACAGGCTGGCTGGAAGCCCTGCACAGT	206					
1365	agacagctgttcagaaggagaccaccaactcagaatgaataataatgaagaaggagaagcta	1364					
205	GGACACACTGCCCCGAGGGCGCCCCACAGGGCAGGGGAAAAAGGACGACACAGAGGACTA	146					
1365	tgatatgatcttactcagatcttcgaatgatgcatgtgctcaggttgattgatgacaactctcc	1424					
145	TSATTA-----TCTGGACTGGGAATATGAGACTGTCTTACCTCGGGACCCCGCATGCCGGCC-	91					
1425	ttccttattccaactatgctcgaagttgccaagaagatccctaaacttggtgatactcat	1484					
92	--CCTTCTCTCTCAGTGGCCCTCAGCAGCCACAGACACCCTCAAGACCTGGTGCTCACTACAT	35					
1485	tgctctgaagaaggagacttggaactatgcgcc	1517					
34	CGCAGCAGAGGAGGGGAGCTGGAGCTACGCGCC	2					
BF420450	484 bp	mRNA	EST	28-NOV-2000			
U1-R-BJ2-bpw-f-03-0-U1.s1	U1-R-BJ2	Rattus norvegicus	cdna clone				
U1-R-BJ2-bpw-f-03-0-U1.3			mRNA sequence.				
BF420450							
BF420450.1	GI:11408479						
	EST.						
	Norway rat.						
	Rattus norvegicus						
	Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;						
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;						
	Rattus.						
	1 (bases 1 to 484)						
	Bonaldo,M.F., Lennon,G. and Soares,M.B.						
	Normalization and subtraction: two approaches to facilitate gene						
	discovery						
	Genome Res. 6 (9), 791-806 (1996)						
	97044477						
	Contact: Soares, MB						
	Program for Rat Gene Discovery and Mapping						
	University of Iowa						
	451 Eckstein Medical Research Building Iowa City, IA 52242, USA						
	Tel: 319 335 8250						
	Fax: 319 335 9565						
	Email: msoares@blue.weeg.uiowa.edu						
	Oligo-dT track not found. Not a site shown in beginning of sequen						
	is likely internal to the message. cdna library Preparation: M.B.						
	Scores lab Clone distribution: clones will be available through						
	Research Genetics (www.resgen.com)						
	Seq primer: M13 Forward						
	POLYA-No.						

```

FEATURES
SOURCE
    Location/Qualifiers
1. .484
   /organism="Rattus norvegicus"
   /strain="Sprague-Dawley"
   /db_xref="taxon:10116"
   /clone="UI-R-BJ2-bpw-f-03-0-UI"
   /clone_lib="UI-R-BJ2"
   /lab_host="DH10B (Life Technologies)"
   /note="Vector: pUT3D-Pac (Pharmacia) with a modified
polylinker. Site_1: Not I; Site_2: Eco RI; The UI-R-BJ2
library is a subtracted library derived from the following
tissues: heart, atrium at 15 dpc, ventricle at 16.5 dpc,
atrium at 16.5 dpc, ventricle at 13 dpc, ventricle at 15
dpc, AV canal at 15 dpc. For a detailed description of
the library from which this clone was derived, please
visit our web site at ratest.eng.uiowa.edu. The

```

```

subtraction has been previously described in (Bonaldo,
Lennon and Soares, Genome Research 6:791-806, 1996)
TAG_SEQ=None found"
BASE COUNT      98 a      137 c      164 g      85 t
ORIGIN

```

Query Match	5.08;	Score 236.2;	DB 148;	Length 484;
Best Local Similarity	68.28;	Pred. No. 4.5e-57;		
Matches 328; Conservative	0;	Mismatches 153;	Indels 0;	Gaps 0

Oy	1522	gtctctgcgcccccgtgcacaaatctataaagcattcttaaacatgctcaatgcccctcaagcg	1581
Dd	4	gccgccacccttgcagaaagcaggagacctctccgaacacggctacctggggcgctggccctggccgc	63
Oy	1582	atctgtaggaagtacaaaanaagtcgatttctatgcatcacasgatgtaaacctttaagact	1641
Dd	64	atccgcacgcaaaatpataagaaagttcagagtttcgtggcctacacgacgcacatttccacagc	123
Oy	1642	cgtaaacgtattatgcacatgatataaggaatcttggaaccttaactctatgagggaaattgga	1701
Dd	124	cgttaggcccaaccggcgtgaacctggggctccttagggccactgctctacgggagagtgcgga	183
Oy	1702	gacacactgtgtatattatattaaagatcaagcaagcagacacataaacatcttaccctcac	1761
Dd	184	gacttcactgttgattgtgtttcaagaacccgacgacggcggtatcaaacattccacccttcatt	243
Oy	1762	ggaatacctgattgtccgtctctctgttatcttaagagagattccaaaagtggtaaacatttg	1821
Dd	244	ggagatccgtatgcagttggcgctgtgcatagcagggagcgctccacgaggggtatagacactg	303
Oy	1822	aagatttcccaattctgtccagaggaagaataattcaatatlaaatgacagtgactgttagaa	1881
Dd	304	aagattttggccgattcccccagagagagactgttcaattcagctggagacgcgtaccgccagag	363
Oy	1882	gattggccaactaatatgaagtcctgtgctgcacacggctattactctagtcttgtaatt	1941
Dd	364	gatggcgcttcgaagctgtggatgcgcgggtgtgtataccgcgtactatgtccagacgccttgat	423
Oy	1942	atgtagagagatctagctttaaagactcaatctgtgcccctctctatctgctacaaagaactc	2001
Dd	424	cccgagagagacactggtctccggcgctgattggcccccctctcatctgctgttmaagaattct	483
Oy	2002	g 2002	
Dd	484	G 484	

RESULT	LOCUS	DEFINITION	BE847128	401 bp	mrna	EST	26-SEP-2000
BE847128	LOCUS	DEFINITION	BE847128	401 bp	mrna	EST	26-SEP-2000
			uw22b1.y1		Soares mouse 3NDMS Mus musculus cDNA clone IMAGE:3417405		
					similar to gb:U05573 Mus domesticus coagulation factor VIII mRNA		
					, complete (MOUSE);, mRNA sequence.		

ACCESSION	BE847128	
VERSION	BE847128.1	GI:10305395
KEYWORDS	EST.	
SOURCE	house mouse.	
ORGANISM	Mus musculus	

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus  
AUTHORS 1 (Pages 1 to 401)  
TITLE NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>  
JOURNAL National Cancer Institute, Cancer Genome Anatomy Project (CCAP), Tumor Gene Index  
COMMENT Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.

This clone is available royalty-free through LNL; contact the  
 IMAGE Consortium ([info@image.lnl.gov](mailto:info@image.lnl.gov)) for further information.  
 MGI:1093217  
 Seq primer: -40RP from Gibco.  
 Location/Qualifiers











GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2001, 13:22:16 ; Search time 41.69 Seconds  
(without alignments)  
4668.282 Million cell updates/sec

Title: US-09-689-430-2

Perfect score: 7869

Sequence: 1 M0ELSTCFELCLRFCEFS.....WHQIALRMVELGCEAQLDLY 1471

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
1: SPREMBL\_16:\*  
2: sp\_archaea:\*  
3: sp\_bacteria:\*  
4: sp\_fungi:\*  
5: sp\_human:\*  
6: sp\_invertebrate:\*  
7: sp\_mhc:\*  
8: sp\_mhc:\*  
9: sp\_mhc:\*  
10: sp\_mhc:\*  
11: sp\_mhc:\*  
12: sp\_mhc:\*  
13: sp\_mhc:\*  
14: sp\_virus:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6401	81.3	2343	6	018806
2	6360	80.8	2343	6	062730
3	2402.5	30.5	2224	4	043737
4	2394.5	30.4	2183	11	088783
5	2371.5	30.1	2258	6	096LPI
6	1792	22.8	1157	11	092024
7	1703	21.6	1084	11	09JL97
8	1619.5	20.6	1048	6	09XJ27
9	1334	17.0	891	4	075180
10	1214	15.4	782	4	075659
11	1104	14.0	216	4	014286
12	664	8.4	463	11	09R1X9
13	662.5	8.4	480	4	043854
14	661	8.3	426	11	09WTS3
15	655.5	8.3	480	11	035474
16	514.5	6.5	363	6	077718
17	469.5	6.0	901	4	09H2E4
18	469.5	6.0	901	4	09H2D5
19	469.5	6.0	906	4	09H2E3

20	469.5	6.0	906	4	09H2D4	09H2d4 homo sapien
21	450.5	5.7	921	11	090X38	09q38 rattus norv
22	429.5	5.5	704	4	09H2EL	09h2el homo sapien
23	429.5	5.5	779	4	09NTR3	09ntt3 homo sapien
24	411.5	5.2	858	5	076470	076470 lytechinus
25	383.5	4.9	555	4	09H2E2	09h2e2 homo sapien
26	273	3.5	51	4	09U005	09uq05 homo sapien
27	271	3.4	1128	11	088442	088442 mus musculu
28	266	3.4	764	11	054860	054860 mus musculu
29	262	3.3	1158	4	014113	014113 homo sapien
30	251	3.2	3843	5	09U5D0	09u5d0 drosophila
31	251	3.2	3843	5	09VU94	09vu94 drosophila
32	247	3.1	224	11	09R1M6	09r1m6 mus musculu
33	223	2.8	722	11	092100	092100 mus musculu
34	223	2.8	728	6	097567	097567 bos taurus
35	221	2.8	719	11	061281	061281 mus musculu
36	217	2.8	734	4	09NDB5	09ndb5 homo sapien
37	217	2.8	1331	4	09UHC6	09uhc6 homo sapien
38	217	2.8	1339	4	09U012	09u012 homo sapien
39	211	2.7	192	11	090WQ1	090wq1 mus musculu
40	211	2.7	1283	5	094887	094887 drosophila
41	211	2.7	1284	5	09VTU5	09vtu5 drosophila
42	207.5	2.6	737	5	017305	017305 caenorhabdi
43	206.5	2.6	737	5	018163	018163 caenorhabdi
44	185.5	2.4	911	11	035407	035407 mus musculu
45	185	2.4	3396	5	09VMS5	09vms5 drosophila

## ALIGNMENTS

## RESULT 1

018806 ID 018806 PRELIMINARY; PRT; 2343 AA.  
AC 018806;  
DT 01-JAN-1998 (TREMBLrel. 05, Created)  
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
DE FACTOR VIII.  
GN F8.  
OS Canis familiaris (Dog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
OX NCBI\_TaxID=9615;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=LIVER;  
RA Cameron C., Notley C., Hoyle S., McGlynn L., Hough C., Kamisue S.,  
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF016234; AAB7412.1; -.  
DR HSSP; P00451; ICFG.  
DR InterPro; IPR001117; -.  
DR Pfam; PF00394; Cu-oxidase; 3.  
DR PROSITE; PS01285; FA58C\_1; 2.  
DR PROSITE; PS01286; FA58C\_2; 2.  
DR PROSITE; PS00079; MULTICOPPER\_OXIDASE1; 3.  
DR SMART; SM00231; FA58C; 1.  
SQ SEQUENCE 2343 AA; 265829 MW; A854FAE571C3B399 CRC64;

## Query Match

Best Local Similarity 81.3%; Score 6401; DB 6; Length 2343;  
Matches 1268; Conservative 85; Mismatches 108; Indels 886; Gaps 4;

QY 1 M0ELSTCFELCLRFCEFSATRRYVIGAVELSMYMSD-LGELPYARPPRPYKSPF 59  
DB 1 M0ELSTCFELCLRFCEFSATRRYVIGAVELSMYMSDLSALHADPSFSSRVGSLPL 60  
QY 60 NTSVYKKTLFEFTVHLFNIAKRPMPMGILGPTIQAEVYDVVITLKNASHPVSLHA 119



Db	61	TISVYRTKTVEEFDDLEFNIAKRRPPMMGLLPITQAEVYDVTYIVLKNASHVLSHA	120
Qy	120	VGVSYWKASBEGAEXDDQTSQREKEDDKVYPCGSHTYVMQVLKENGPMASDPICLTYSLS	179
Db	121	VGVSYWKASBEGAEXEDQTSQREKEDDKVYPCGSHTYVMQVLKENGPMASDPICLTYSLS	180
Qy	180	HVDLYKDLNSGLIGALLYCRGSLAKKERTQTLAKTILLFAVDEKSMHSEKNSLMODR	239
Db	181	HVDLYKDLNSGLIGALLYCRGSLAKKERTQTLAKTILLFAVDEKSMHSEKNSLMODR	238
Qy	240	DAASRAMPKMHVGVYVNSRSLPGICGHKRSYVNHVIGMGTPEVHSIFLEGHTFLVRN	299
Db	239	---AEAGHELHTINGYVYVNSRSLPGITLGHKRSYVNHVIGMGTPEVHSIFLEGHTFLVRN	294
Qy	300	HRQASLEISPTFTLAQTLMDLGOFLFCHTISHQHDMEAYVYKVDSCPEEPOLRMKN	359
Db	295	HRQASLEISPTFTLAQTLMDLGOFLFCHTISHQHDMEAYVYKVDSCPEEPOLRMKN	354
Qy	360	EEADYDDDLDSMDVYRPFDDNSPSTQIRSVAKKHPRKVMVNHIAAEEEDMDYAPLV	419
Db	355	ED-KDYDGLYDSMDVYRPFDDNSPSTQIRSVAKKHPRKVMVNHIAAEEEDMDYAPSGP	413
Qy	420	APDRSYKSOYLNNQPORIGRKKYKVRMAATDEFKFKREALIOHSGILGPLLYEEVDY	479
Db	414	TPNDRSHKNLYLNNQPORIGKKKYKVRMAATDEFKFKREALIOHSGILGPLLYEEVDY	473
Qy	480	LLIIFKNQASRPYNIYPGHTDVRPLYSRRLPKGVKHLKDPILPGLIIFYKWTYVBDG	539
Db	474	LLIIFKNQASRPYNIYPGHTDVRPLYSRRLPKGVKHLKDPILPGLIIFYKWTYVBDG	533
Qy	540	PTKSDPRLTTRYSSFFVMEKRDLASGLIGPLLTCKESVDORGNOMSDKRNVLFSVD	599
Db	534	PTKSDPRLTTRYSSFFVMEKRDLASGLIGPLLTCKESVDORGNOMSDKRNVLFSVD	593
Qy	600	ENRSGVYLENIOFLPNPAGVOLDEPEFOASIMHSINGVYFDSIQLSVCLHEVAYWYIL	659
Db	594	ENRSGVYLENIOFLPNPAGVOLDEPEFOASIMHSINGVYFDSIQLSVCLHEVAYWYIL	653
Qy	660	SIGAQDFLSYFFSGYTPKHKMYVEDTLTLPFSGSETFVMSMENPGYLILGCHNSDPNR	719
Db	654	SVGAQDFLSYFFSGYTPKHKMYVEDTLTLPFSGSETFVMSMENPGYLILGCHNSDPNR	713
Qy	720	GMTALLKVSQCDKNTGQYEDSYEDISAYLLSKNNAIEPRFSQNSRHPSTRQOKFNT-	778
Db	714	GMTALLKVSQCDKNTGQYEDSYEDISAYLLSKNNAIEPRFSQNSRHPSTRQOKFNT	773
Qy	779	---	778
Db	774	TPENDIEKIDLOSGERTOLIAQSVSSDDLMLLQONPTPRGLISDLREATDRADHSR	833
Qy	779	---	778
Db	834	GAIERNKGPRPVASLRLPRLRHSDBRETPPEBLORLNENIGTNTTVBLKLLDLKISSS	893
Qy	779	---	778
Db	894	DSLMTSPTIPSDKLAATEKTGSLGCPMMSVHFNSHLCTIVFGNNSHLLIOGVLELSE	953
Qy	779	---	778
Db	954	EDNDSKLLIADPLMNIQESSLRENNVLSMESNRLFKBEIRIJPASLIKONALFKVNISYKT	1013
Qy	779	---	778
Db	1014	NRAVNLTLTNKKTVAIPTLLIENSTSWQOIMLEBNTFEEKEVYLSLHNETFMDRNTAL	1073
Qy	779	---	778
Db	1074	GLNHSVNKTTLSKNVEMAHQKEDBPVPLRAENPDLSSKIFPLPDWIKTHGKNSLSSQR	1133
Qy	779	---	778
Db	1134	PSPKOLTSLSGSEKSVKQDNLFSEKVVVGDEFTKDTLOEIFPNNKSIFEPANLANVOEN	1193

OY	779	----	778
Db	1194	DTYNQEKKSPEELERKEKELTOENVALPOAHITMIGTKNPLKNEFLSTKQNVAGLEEDPYT	1253
OY	779	----	778
Db	1254	PILODTSLANDSPHSEIITHMANFSKIREBANLEGLOTNOMVEREPSTTRMSSNASOH	1313
OY	779	----	778
Db	1314	ITQGRKSLKQPRLSQGEIIFERKVIYLANDSTQWSKNMNYLQAGTTLQLEYNEKKRAIT	1373
OY	779	----	778
Db	1374	QSPLSDCSMRNHVTIOMNDSALPVAKESASPVYRHTDLTKIPSOHNSHLPASACNTYFR	1433
OY	779	----	778
Db	1434	ERTSGVOBGSHFLOEAKRNMLSLAEVYTLGITBGQKFESSLGASATNOPMYKKELENTVLLQ	1493
OY	779	----	778
Db	1494	PGLSEPTDXKVELLSQVHVNDDEDSFPKTKTSNDSRGHLDLMGKLFLOKQTPYKMKNTSPG	1553
OY	779	----	778
Db	1554	KVPELKMATESSEKIPSKLLGLVAMDNIHQDTQIPSEMKSQKSQTNATFKKDKTILPLG	1613
OY	779	-----PVLKRNORETRITTLTSDOEI	801
Db	1614	PCENNDSSTAINEGODKQREAMAMAKQBERGLCSQNPVSNHNRREITVTTLOPEDEKF	1673
OY	802	DYDDTIIVEMKKEDFDIYDEDEDNQSPRSQKTIHNFYIAAVERLMDYGMSSPHVLRNRA	861
Db	1674	EYDTEFTEMKREDFDIYGYEMQGLRSQOKTRHFTIAAVERLMDYGMSSPHLRNRA	1733
OY	862	QSGSVPQFKKVVFOEFTDGSFTQPLYRGELNEHLGLBPYIAAEVEDNIWTFPRNQASRP	921
Db	1734	QSGVQOQFKKVVFOEFTDGSFTQPLYRGELNEHLGLBPYIAAEVEDNIWTFPKNQASRP	1793
OY	922	YSFSSLSLISVEEDROGAEFRKRFVVRNETKTYFMKVQONHMAPTDEDEDCAMAYFSVD	981
Db	1794	YSFSSLSLISYDEBGOGAERFRKRFVVRNETKTYFMKVQONHMAPTDEDEDCAMAYFSVD	1853
OY	982	LEKOVHSGLLGPRLVCHTNTPLNPAHQROYVDFALFTIPDETQSYVFTEMERNCRAP	1041
Db	1854	LEKOVHSGLLGPRLVCHTNTPLNPAHQROYVDFALFTIPDETQSYVFTEMERNCRAP	1913
OY	1042	CNIOMEDPTEKENYRFHAIINGYIMDTPLPGLVMAODQRIWYLLSGNSMENIHSIHFSGHV	1101
Db	1914	CNYOKEDPTLKENRFHAIINGYIKDTPLPGLVMAODQKVBYWYLLSGNSMENIHSIHFSGHV	1973
OY	1102	FTVAKKEEYKALYLNLYPGVFEIVBEMLPKSAQIMWRECLIGHILAGNSTPLFYUYSKQO	1161
Db	1974	FTVAKKEEYKMAVYNLYPGVFEIVBEMLPKSAQIMWRECLIGHILAGNSTPLFYUYSKQO	2033
OY	1162	TPGLMASGHIRDQOITASAGYQGMARKLRLHYSGSINAMSTKEPFSYITKDLLAPMIIH	1221
Db	2034	TPGLMASGHIRDQOITASAGYQGMARKLRLHYSGSINAMSTKEPFSYITKDLLAPMIIH	2093
OY	1222	GIRTOGAROKFESSLYISOPLIMSLDGBKWQTYRGNSTTLLWVFPGVNDSSGIRHNHFNFP	1281
Db	2094	GIMTOGAROKFESSLYISOPLIMSLDGNKMHYSRGNSTGTLWVFPGVNDSSGIRHNHFNFP	2153
OY	1282	PIIARVIRLHPHYSTRSTRLMELMGCDLNSGSMPLGMEKSAISDAOITASSYTNMFPAT	1341
Db	2154	PIIARVIRLHPHYSTRSTRLMELMGCDLNSGSMPLGMEKSAISDAOITASSYTNMFPAT	2213
OY	1342	WSPSKARLHLOGSSNMNRQVNNPKEMLOVDQOKIMKUYLGVTQGVKSLLSMIVKFEFLI	1401
Db	2214	WSPSOARLHLOGGRNMMRPPANNPKEMLOVDPRKIMKUYLGVTQGVKSLLSMIVKFEFLI	2273



```
QY 1402 SSSODGHWLTFEFGNGKVVFEQGNDSFTPVVNSLDPLLTFRYLRIHPQSVWHTALRME 1461
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2274 SSSODGHWMTLFLQNGKVVFEQGNDSFTPVVNSLDPLLTFRYLRIHPQSVWHTALRME 2333
QY 1462 VUGCEAQ 1468
      |||||
Db 2334 VUGCDTQ 2340

RESULT 2
062730 PRELIMINARY: PRT: 2343 AA.
ID 062730:
AC 062730:
DT 01-AUG-1998 (TREMBLrel. 07, Created)
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE FACTOR VIII.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_Taxid=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=KIDNEY, SPLEEN;
RA Gordy P.W., Bowen R.A.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF049489; AAC05384.1; -.
DR HSBP: P0451; ICRG.
DR InterPro: IPR00421; -.
DR InterPro: IPR00117; -.
DR Pfam: PF00394; Cu-oxidase; 3.
DR Pfam: PF00754; F5_P8_type_C; 2.
DR PROSITE: PS01285; FA58C_1; 2.
DR PROSITE: PS00786; FA58C_2; 2.
DR PROSITE: PS00079; MULTICOPPER_OXIDASE1; 3.
DR SMART: SM00231; FA58C; 1.
DR SMART: SM00231; FA58C; 1.
SQ SEQUENCE 2343 AA; 265613 MW; F612D744ADADD99 CRC64;

Query Match 80.8%; Score 6360; DB 6; Length 2343;
Best Local Similarity 53.7%; Pred. No. 0;
Matches 161; Conservative 88; Mismatches 112; Indels 886; Gaps 4;

QY 1 MOEISTCFPLCLFRFCFSATRRYLLGAVELSDMYMOSD-IGELPYDAFRPPRVKSPFE 59
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 MOVELYTCFCFLPLPPSISATRKTYLGAVELSDMYMOSDLSALHADVSFSSRVGSLPL 60
QY 60 MTSVYKKTLEFVEFTVHLFENIAKPRPPMGLIGPTIOAEVYDTVVITLKNMASHVSLHA 119
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 TTSVYRKTVFVEFTDLEFNIAKPRPPMGLIGPTIOAEVYDTVVITLKNMASHVSLHA 120
QY 120 VGVSVYKASSEGAEYDDQTSQREKEDDKVFPGGSHYVWQVLYKENGPMASDPLCTLYSYLS 179
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 121 VGVSVYKASSEGAEYDDQTSQREKEDDKVFPGGSHYVWQVLYKENGPMASDPLCTLYSYLS 180
QY 180 HVDLYKDNSGLIGLALVCRSGSLAKEKTKOTLHFKILLFAVDEKSKWHSFKNSLMDR 239
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 181 HVDLYKDNSGLIGLALVCRSGSLAKEKTKOTLHFKILLFAVDEKSKWHSFKNSLMDR 238
QY 240 DAASRAVPMKHTVNGVYVNSRLPGLIGKRSVYVWVIGMTTPVHSTIFLEGHTFLVFN 299
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 239 -----AEAGHELTHTINGVYVNSRLPGLIGKRSVYVWVIGMTTPVHSTIFLEGHTFLVFN 294
QY 300 HROASLEISPTFLTAQTLTMDLGFLLFCHISSHQHDGMEAYVYKVDSCPEEPQLRMKN 359
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 295 HROASLEISPTFLTAQTLTMDLGFLLFCHISSHQHDGMEAYVYKVDSCPEEPQLRMKN 354
QY 360 BEAEEDYDDLTSEMDVYVRFDDNSPSFQIRSVAKKHKHTKVHNTAAEEDMDVAPLVL 419
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 355 ED-KYTDGSLGSDMDVYVRFDDNSPSFQIRSVAKKHKHTKVHNTAAEEDMDVAPLVL 413
QY 420 ADDDSYKSOYLNNGPORIGRKYKRVFMAVYDDEFKTEALIOHSGSLIGPLLYGEVGT 479
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
Db 414 TPNDRSKKNLVLNNGPORIGRKYKRVFMAVYDDEFKTEALIOHSGSLIGPLLYGEVGT 473
QY 480 LLITRNQASRPYNYNIPPHGITYVRLPRLPKGYKHLKDFILILGEIFKYYKWTYVVEG 539
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 474 LLITRNQASRPYNYNIPPHGITYVRLPRLPKGYKHLKDFILILGEIFKYYKWTYVVEG 533
QY 540 PTKSDPRLTRYSSFFVNNERDLASGLIGPLILCYKESVDQGNQIMSDKRWILFESVD 599
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 534 PTKSDPRLTRYSSFFVNNERDLASGLIGPLILCYKESVDQGNQIMSDKRWILFESVD 593
QY 600 ENRSWYLTENIGFLPNPAGVQLEDEPFGASIMHISINGVDSQLSVCLHEVAYWYL 659
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 594 ENRSWYLTEDMORFLPNADVQPHDEPQLSNIMHSINGVDFDNLQSVCLHEVAYWYL 653
QY 660 SIGAOTDFLSVFSFGYTFKHKWYEDTLTLPFPGSGTYVMSMENGLMGLGHSNDFRR 719
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 654 SVGAOTDFLSVFSFGYTFKHKWYEDTLTLPFPGSGTYVMSMENGLMGLGHSNDFRR 713
QY 720 GMTALLKVSCKNNGDYEDSYEDISAYLLSKNNAIEPRFSQNSRHPSTROKOFNAT- 778
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 714 GMTALLKVSCKNNDIYEDYEDIPPLLENVNIKRFSQNSRHPSTROKOLKATF 773
QY 779 ----- 778
Db 774 TPENDIEKIDQSGERTOLIKASVSSDLMGLIGNPTPRLPLSDLEATDRADDSHR 833
QY 779 ----- 778
Db 834 GAIERNKGPPEVASLRPELRHSEDEFTPEPELOLRLENIGTNTVELKLDLKISSSS 893
QY 779 ----- 778
Db 894 DSLMTSPITPSDKLAATEKTSIGPPMNSVYFNHGLGTYVGNSSHLQSGVPLELSE 953
QY 779 ----- 778
Db 954 EDNDSKLEAPLMNIOESSLRBNVLSMESNRLEFKEBIRGPASLIKDNALFKVNISSYKT 1013
QY 779 ----- 778
Db 1014 NRPVNLTTNRRKRAVAPTLLENSTSWQDIMELEMRTEKFEVTSLIHNETFMDRNTAL 1073
QY 779 ----- 778
Db 1074 GLNHNKNTTLESKNVEMAHOKKEDVPPLRAENPDLSSSKIPPLPWIKTHGKNSLSSEOR 1133
QY 779 ----- 778
Db 1134 PSPQUTLSGSEKSVKQDNFLSEKVVYVGEDEFKTDTELQEIFPNKNSIFFANLANVOEN 1193
QY 779 ----- 778
Db 1194 DTYNOEKSLIEIERKEKLTQENVALPOAHMTIGTKNPLKNLFLSTKONVAGLEOPYT 1253
QY 779 ----- 778
Db 1254 PTLQDTRSLNDSPHSEGTHMANFSKIREBANLEGLGNQTNQWVERPSTTRMSSNASHV 1313
QY 779 ----- 778
Db 1314 ITQGRKRSILKQRLSGELKFERKYIANDTSTQWKNMNYLAQGLTLQIEVNEKERRAIT 1373
QY 779 ----- 778
Db 1374 OSPLSDCSMRNHVTTIQMNSALPVAKESAPSVRHTDLTKIPSOHNSHLPASACNYFER 1433
QY 779 ----- 778
Db 1434 ERTSGVQBSHPLQDAKRNNSLAEVTLGITBEGCKFSSLGKSAATNOPMYKKELVTLQ 1493
QY 779 ----- 778
Db 1494 POLSETSDKVELLSQVNHQDEDSFTTKTSNDSRGLHLMGKIFLQKTKGQPVKMNKNTNSPG 1553
```



QY 779 ----- 778  
 Db 1554 KVPFLKWTATESSEKIPSKLLGLVAMDNHYDTQIPSEEMKSOQKSTNTAFKRKDTILPLG 1613  
 QY 779 -----PVLKRNORRITITTTLOSDEEI 801  
 Db 1614 PCENNDSTAINEGODKPOREAMMAKOGEPRLCSONPVPKSHORREITVTTTLOPEEDKF 1673  
 QY 802 DVEDDTISVEMKKEPDIDYDEDEMOSSPRFOKTRHYFIATVERLMDVGMSSPVLKRNRA 861  
 Db 1674 EYDDTFSIEKREDFDITGDEYEDOGCSFOKTRHFIATVERLMDVGMSSPVLKRNRA 1733  
 QY 862 OSGSVPOFKVYVOEFTDSEFTOPLYRGELNEHLGLGPYRAEVEDNIMVTERNQSAR 921  
 Db 1734 OSGDVQOFAKVVQFETDSEFTOPLYRGELNEHLGLGPYRAEVEDNIMVTERNQSAR 1793  
 QY 922 YSFYSSLLIYEEOGCAERKRVNPNKTKTFWKVQHMAATPKDEFDCKANAYESD 981  
 Db 1794 YSFYSSLLIYDEDEGCAERKRVNPNKTKTFWKVQHMAATPKDEFDCKANAYESD 1853  
 QY 982 LEKDVHSGILGPLVCHTNTLNPAGHROVTVQEFALFTTJFDETKSVYFTENNERCRAP 1041  
 Db 1854 LEKDVHSGILGPLVCHTNTLNPAGHROVTVQEFALFTTJFDETKSVYFTENNERCRAP 1913  
 QY 1042 CNIQMEDPTFKENYRPHAINGYIMDTPLGLVMAODORIRWYLLSMGSNENIHSTHESGHV 1101  
 Db 1914 CNVQKEDPTLKEFRPHAINGYIMDTPLGLVMAODORIRWYLLSMGSNENIHSTHESGHV 1973  
 QY 1102 FVVRKKEEKMAKXNLYPGVFEVEMLPKAGIWRVECLIGENLHAGMSTLFLVYSKQC 1161  
 Db 1974 FVVRKKEEKMAKXNLYPGVFEVEMLPKAGIWRVECLIGENLHAGMSTLFLVYSKQC 2033  
 QY 1162 TPLGMAAGHIRFOLITASGOYOMAPKLARLHSGSINAMSTKEPFSIKYVDLLAPIT 1221  
 Db 2034 TPLGMAAGHIRFOLITASGOYOMAPKLARLHSGSINAMSTKEPFSIKYVDLLAPIT 2093  
 QY 1222 GIKTGOAROKFSSLYISOFTIMYSLDGKWKOTYRGNSTGLMWFGVNDSSGKHNIFNP 1281  
 Db 2094 GIMTGOAROKFSSLYISOFTIMYSLDGKWKOTYRGNSTGLMWFGVNDSSGKHNIFNP 2153  
 QY 1282 PIIARIYIRLHPHYSTRSLRMLMGLCDLNSCMLGMEKASIDAOITASSYFTNNFAT 1341  
 Db 2154 PIIARIYIRLHPHYSTRSLRMLMGLCDLNSCMLGMEKASIDAOITASSYFTNNFAT 2213  
 QY 1342 WSPSKARLHLOGRSNAMPQOVNPKEMLOYDFQKTKKVTGVTTOGVSLTSMYVKEFL 1401  
 Db 2214 WSPSKARLHLOGRSNAMPQOVNPKEMLOYDFQKTKKVTGVTTOGVSLTSMYVKEFL 2273  
 QY 1402 SSSQDGHOWTLEFFONGKRVFQGNODSFTPVVNSLDPELLTTRYLRILHPQSMVHQIALRME 1461  
 Db 2274 SSSQDGHOWTLEFFONGKRVFQGNODSFTPVVNSLDPELLTTRYLRILHPQSMVHQIALRME 2333  
 QY 1462 VLGCEAQ 1468  
 Db 2334 VLGCDTQ 2340

RESULT 3  
 043737 PRELIMINARY: PRT: 2224 AA.  
 AC 043737  
 DT 01-JUN-1998 (TREMBlrel. 06, Created)  
 DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
 DE FACTOR V.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.  
 NCBI\_Taxid=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Bird C.;

RL Submitted (JAN-1998) to the EMBL/GenBank/DBU databases.  
 DR EMBL: 299572; CAB16748.1; -  
 DR HSSP: P00450; 1KCW.  
 DR InterPro: IPR000421; -  
 DR InterPro: IPR001117; -  
 DR Pfam: PF00394; Cu-oxidase; 3.  
 DR Pfam: PF00754; F5\_F8\_type-C; 2.  
 DR PROSITE: PS01285; FA58C\_1; 2.  
 DR PROSITE: PS01286; FA58C\_2; 2.  
 DR PROSITE: PS00079; MULTICOPPER\_OXIDASEL; 2.  
 DR SMART: SM00231; FA58C; 1.  
 DR Blood coagulation.  
 KW SEQUENCE 2224 AA; 251673 MW; 1013108049690EAB CRC64;

Query Match 30.5%; Score 2402.5; DB 4; Length 2224;  
 Best Local Similarity 26.3%; Pred. No. 8-177;  
 Matches 598; Conservative 280; Mismatches 489; Indels 909; Gaps 37;

QY 22 RRYVLGAVELSDWYMOISDLGELPVDARFPPRVKSPFPNTSVY-YKRTLFVEFTVHLFNI 80  
 Db 32 RQFYVAAGISWSVRPE-----PTNSLNLSTVSFKKIYREVEPY-FKK 75  
 QY 81 AKRPPPMGLSPTIOAEVYDVTYITLKMAHSIPVSLHAGVSYKASGAEYDOTSOR 140  
 Db 76 EKPOSTISGLGLPTLYAEVDIIVHEKNAKADPLSIHQGLRYSKLSGASYLDTHTPPA 135  
 QY 141 EKEDKVFPGSGSHYVQVLEKENGPMASDPLCTLYSLSHVDLVKRLNGLGALLVCE 200  
 Db 136 EKMDDAVAPRETTEKSTIEDSGPTHDPCLTHIYSHENLIEFNGLGLPLICK 195  
 QY 201 GSLAKERTQ-TLHK-FLILFAVDEKSWHSETRKSLMODRDAASAMPKMYNGYV 258  
 Db 196 GTLLEGCTQTFPKQIYLLFAVDEKSMQSSS-----LMTYGVYV 239  
 QY 259 RSLPGLIGCRKSYVHWVAVGMTPPEVHSIFLEGHFLVARNHQALESPTFLAQT 318  
 Db 240 GTMPDITVCAHDIISWLLGMSGPELFSIFNGQVLEONHHRVSAITVLSATSTANNT 299  
 QY 319 IMLDGOFLCHTSHSHQHDGMEAYVQVDSCEPRLRMKNNEAEYDDDLTDEMDYVR 378  
 Db 300 VPEEGKWIISLTPKHLQAGMAYIDIKKCPKTRMLKITTRE----- 342  
 QY 379 FDDONSPSFTQIRSVAKHKRTVHYIAAEEEDMDYAPLVLPDDRYSKQYLANNGPOT 438  
 Db 343 -----QRBHKKRWEYFLAAEVIWDAVYIPANMOKKYSQHLDNFSNY 387  
 QY 439 GRKYKVRMAATYDETFKTRREAL---QHESGILGPLLYEGVDTLITIKNOASRPYNT 495  
 Db 388 GKHYKVMYQYDESEF-TKHVTNPNMKEDGILGPIIRAQVDTLKYKKNASRPYSY 446  
 QY 496 PHGIT-----DYRPLYSRLPKGVKHLKDFPLIPGEIRKMYVTVEDGPKSPRCIT 549  
 Db 447 PHGVTFSPYDEVNSFT-----SGRNNYMIKAVQPEFTYTKMNLIFEPPEINDAQCIT 502  
 QY 550 RYVSSFYVMERDIASGLIGPLICYKESVDQGNQIMSDKRVNLLFSVDENRSMYLTEN 609  
 Db 503 RPYSDVDIMRIDIASGLIGLILICKRSIDRQIQRAADIEQAVFAVDEKSKWYDLN 562  
 QY 610 IQRFLEPNPAGVOLEDEPEQASNTMHSINGVPSLT-QLSVCLAEVAYWILSIGQOTDL 668  
 Db 563 INKFCENPDEVKRDDPRFYESNTMSTINGVPSITTLGFCDDTVQMHFCVGTQNEL 622  
 QY 669 SVFSGYTFKHKVAVYDITLFFPSGETYFMSMENGILLICHSN----- 714  
 Db 623 TIFTGHSFYIGKRHEDITLFPMRGESYVTINDVGTMLLSMNSPSKRLKLFADV 682  
 QY 715 -----DFRNR----- 719  
 Db 683 KCIIPDDEDSYEIPEPPESTVMAIRKMDRLPEDESDADVDYQURLAALGIRSFNS 742  
 QY 720 -----GMTAL----- 724



Db 743 SLNOEEEFNTALALENGTEFEVSSNTDIIIVGSNYSPPSNISKFTVNNIAEPQAKPSHQ 802  
QY 725 -----LKVSSCDKNTGDYEDPSYED----- 744  
Db 803 ATTAGSPILRLHIGKNSVLSNSTAEHSFYSIEDPIQPDVGTGIRLSLGAEGFSGOE 862  
QY 745 ----- 744  
Db 863 AKHKGPRVERDOAAKHRSFWMKLLAHKVGRLSDQTGSPSGMRPWEJLPSQDTGSPSBR 922  
QY 745 -----ISAYLSKNN----- 755  
Db 923 PAKDPSPDLLLKOSNSSKILVGRWHLASEKSGYEIIQDTDEDTAVNNMLISQNASRAW 982  
QY 756 ----- 755  
Db 983 GESPPLANKPGKQSGHPKFPVRKSLQVRODQSGKSLKKSQFLIKTKRKKKEKHTHAP 1042  
QY 756 IEPNSF-----SQ 763  
Db 1043 LSPRTFHLRSEAVNTFSERLKLHSLVAKSNTSLPTDLNQTLPMDFGMTASLPDHNQ 1102  
QY 764 NSRH-----PSTR----- 771  
Db 1103 NSSNDTGACSPGLYQTVPEEHQTPPIQDPDMHSTDPHSRSSPELSEMLEYDRS 1162  
QY 772 OKOF-----NATPPYLKRH- 785  
Db 1163 HKSPPTDISQSPSESEHWQTVISPDLSQVTLSPELSQTNLSPDLSHHTTLPQLQNL 1222  
QY 786 ----- 785  
Db 1223 SPALGOMPISPDLSHTTLPDLSHHTTSLDLSTLSDLSQTNLSPELSQTNLSPALGOMPLSPDL 1282  
QY 786 ----- 785  
Db 1283 TTLSLDSQTNLSPELSHMTLSPELSQTNLSPALGOMPISPDLSHTTSLDLSTLSDLSQTNL 1342  
QY 786 -----QREITRTTQSGOEI----- 801  
Db 1343 LSQTNLSPALGOMPLSPDPSHTTSLDLSTLSDLSQTNLSPELSQTNLSPOLSEMPLEADLSQIPL 1402  
QY 802 ----- 801  
Db 1403 TPDLQMTLSPDLGDTLSPRFGMSLSPDLSDQVTLSPDISDTTILPDLQSISPPDLQ 1462  
QY 802 -----DYDDT 806  
Db 1463 IFYSESSQSLILOEFNESPYPDLQMPSPSPPTLNDTFLSKFNPFLVIGLSKGDHY 1522  
QY 807 ISVEKKF-----DPDIYDE-----DENOSP-----RSFOKKTTHY 837  
Db 1523 IEIIPKEVQSSDDYAEIDVYPDDPKTVDRTNINSSRPDNTAAWYLSNNGNRNY 1582  
QY 838 FIAAVERLMDYGSSSPVHLNR--AOSGSVPO--FKVYFOEFTDGSFQPLYGELN 892  
Db 1583 YIAABEISMDY-----SEYQVQRETDIEDSDLDPEDTYKKVFRKYLDDTFTKRDRGGE 1638  
QY 893 EHLGLGFTYIAEVDNIMTFRNOASRPYSFYSLSYE-----EDROGAEPKRN 944  
Db 1639 EHLGLGFTYIAEVDNIMTFRNOASRPYSFYSLSYE-----EDROGAEPKRN 944  
QY 945 FVKRNETFTYFMKVQHNAAPRKDEDCAMAYFGDVLEKXVHSGLLGILLVCHTNTLP 1004  
Db 1699 AVQPSSTTYVWHATENSGRPSGACRAMAYISAVNPEKIHSGLLGILLVCHTNTLP 1004  
QY 1005 AHGROVTOEALFETTFJEDTKSWFTENMERNCRAPCNIMEDPTFEKENYRFAINQY 1064  
Db 1759 DSMNPMQMRREVLLFMFDEKKSWMYEKSSRSMR-----LTSSEMKKSHFHAINMI 1812  
QY 1065 MDTLPLGVLMAODQIRMYLLSMGSENINSHSGHVFYTKKKEYYKALYNLYPGVET 1124  
Db 1813 Y-SLPGLKMYEQEWVRLHLNLIGGSODIHVVHFGQTLLENGNKQHOLGVMPPLPSGPKT 1871

QY 1125 VEMLEPKAGIRVRECLIGEHLHAGMSTLFLVSNKQCPILGMASGHRDFOITASGQYQ 1184  
Db 1872 LEMASKRGMWMLNTEVENORAGQCPFLIMDRQCRPMGLSTIGIISDSIKASEFLGY 1931  
QY 1185 WAPKLARLYHSGSINAWSTKE--PFS--WIKYDLAPMTIHGIKTGAROKFSSLYIS 1238  
Db 1932 WEPLRLANGSGSYNAWEKLAEPASKPWIOVDMQEVIIITGIGOGAKHYLKSCTYT 1991  
QY 1239 QIIMYSLDGGKKQTYKRNSTGLMVFEGVNDSSGIRKININNPILAYIRLHPHYSIR 1298  
Db 1992 EFYVAYSNQIIMQIFKGNSTRNMYFNGNSDASIKENQDPPIVAYYIRISPRAYNR 2051  
QY 1299 STLHMLMGCDLNSGMLGESKASIDQITASSYFTNMAT--WSPSKARLHLOGRGA 1357  
Db 2052 PTLRLDQGEVNGCSTPLGEMNGKIEKQITASSFKKSMGMDWEPRARLNAQGRVNA 2111  
QY 1358 WRPOVNNPKEMLOVDFQTKMVTGVTGQVRSLLTSMYKBEFLISSQDGHQWTLFPONG 1417  
Db 2112 WQAKANNKQWLEIDLKIKKITAITTGCKSLSEMYVKSVTHYSQGVWKPRYLS 2171  
QY 1418 KV--KVFOGNDSTFPVNSLDPPLTRYLRHPOSWVHQALRMEVLCERADLY 1471  
Db 2172 SMVDKIFEGNNTKGHVKNFNPPIISRFRIVIPKTMQSIARLRLFGC---DIY 2224

RESULT 4  
088783 PRELIMINARY; PRT: 2183 AA.  
AC 088783;  
DT 01-NOV-1998 (TREMBlrel. 08, Created)  
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)  
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
DE MERINE COAGULATION FACTOR V.  
GN F5.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Yang T.L., Cui J., Rehumtulla A., Mousaali M., Kaufman R.J.,  
RA Glnsburg D.;  
RT "The structure and function of murine factor V and its inactivation by  
RT protein C.";  
RL Blood 91:0-0(0012).  
RL [2]  
RP SEQUENCE FROM N.A.  
RA Glnsburg D., Yang T.L., Cui J., Yang A.;  
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U52925; AAC93953.1; -.  
DR HSSP; P00450; IKCW.  
DR MGD; MGI:88382; F5.  
DR InterPro; IPR000421; -.  
DR InterPro; IPR001117; -.  
DR Pfam; PF00394; Cu-oxidase; 3.  
DR Pfam; PF00754; F5\_F8\_type\_C; 2.  
DR PROSITE; PS01285; FA58C\_1; 2.  
DR PROSITE; PS01286; FA58C\_2; 2.  
DR PROSITE; PS00079; MULTICOPPER\_OXIDASE1; 2.  
DR SMART; SM00231; FA58C; 1.  
SQ SEQUENCE 2183 AA: 247228 MW: BF0A8AA/23F60317 CRC64;

Query Match 30.4%; Score 2394.5; DB 11; Length 2183;  
Best Local Similarity 26.0%; Pred. No. 3.6e-176;  
Matches 588; Conservative 296; Mismatches 494; Indels 885; Gaps 35;

QY 8 CFFLCFL--KFC-----FSATRRYILGAVELSDMYMSDGLGELPYDARFPFRYK 56  
Db 7 CFFLLVYLTGVTMAWGSHQAEADLRQFYVAAGILWYHDE-----PTD 51  
QY 57 FPFNTSVYVYKKTLEVEFTVHLNIAKPRPPMNGILGPTIQAEEVDTVIVITLKMAHPVS 116



Db 52 PSLSNIPSEFKKIVYREYEOY-FKKEKPRSSNSGLLGPITYAEVGDYIKVHFRNKAKPKLS 110  
QY 117 LHAVSVYWKASEGAEYDQOTSOREREKDKVFPGGSHTYVMVOYLKENGPMASPLCLTYS 176  
Db 111 IHPOGIKTSKSEEGASYADHTFPAERKDDAVAPGEYETTEMVYSEDSGPRPDDPCLTLH 170  
QY 177 YLSHVDLVKDNSGLLGLALVCREGSLAKEKTOTL--HKFILLFAVDEGKSNHSETKNS 234  
Db 171 YYSYENLQDFENSGILGPLLICKKGTLEDGTQKMFDDKOHVLLFAVDESKS----- 222  
QY 235 LMQDDAASARAKPMKHTYNGVNSLPLGLICHRKSVWHTYGMGTTEVHSLFLEGHT 294  
Db 223 -----RSOSPMLTYTNGFVNKMTMPDITYCAHDVSWHLGMSGGEPLSIHNGOY 274  
QY 295 FLVRNHROASLEISPTFLTAOTLLMDLGOFLFCHISSHOHDGMAYVAVKSCPREPOU 354  
Db 275 LEQONHAKSYTVLVSATSTTAMTMSPEGRWIVSSILPRHYQAGMOAYIDIKMCPKTR- 333  
QY 355 RMKNNEAEDYDDDLTDEMDVYVRDNDNSPSFTQIRSAVAKKHPKTMVHYIAAEEEDMDY 414  
Db 334 -----SPK--TLTROQRMRKMEYFIAAEVITMY 362  
QY 415 APVLAPODRSKSOYLNNGPORIGKRYKRYRPMAYTDTF--KTRPAIQHESGLIGPLL 472  
Db 363 APVLPANMDKTIYRSQHLNDFNSNOIGKHYKRYROYEEETFKTDNPSIKOSGILGPVY 422  
QY 473 XGEVGDLLIFKKNQASRPYNIYPHGITDVRPLYSRLKGY-----KHLKDEPLIG 525  
Db 423 RAQVADTLKIYFKNPASRYSIYPHGV-----PSPYDIGNSSSTSSGHTTIRVOYG 476  
QY 526 EIKFKWTVYEDGPTKSDPCLTRYSSSFVNMERDLAGLIGPLLICRYKESVQDQNGOI 585  
Db 477 EFTYKWMILDERDEPTENDAOCLTRPYSDVOTRIDIASGLIGLLICKSRSLDQGVOR 536  
QY 586 MSDRNVILFSYFDENRSYLTENIQRLPNPAGVLEPPEOASIMHSGYVDSIQ 645  
Db 537 VADIEQAVFAVEFENKSYIEDINIKFENDEKVRDOPKEYSIMSTINGYVESIS 596  
QY 646 -LSVCLHVAWYILSIGAOTDFLSVFESGYEFKHKMYEEDTLTPPFGSEYFMSMNP 704  
Db 597 TLGCFDFTQVQMHFGSVGHDILITIHFGHSFIYGRHEDTLTPPMGESYTYMDVY 656  
QY 705 GLWILGCHNSDFRNRGWTALLKVSSCDKNTGY-YEDSYE----- 743  
Db 657 GTWMLTJNNSPKRRNLRLRFRDYKCNR--DYDNEDSYEYEPAPTSMTRIRIDSLE 713  
QY 744 -----DISAYLLS-----KNAIEPR----- 759  
Db 714 NEFGIDNEDDDYQYLLASSLAGIRSFKNSSLNPENEFNLTALALENSSEFISPDTRVVD 773  
QY 760 ----- 759  
Db 774 SNSSRLSKIIINNKLQFORTLPGSGAIVAGTLNLNLGLDENFVLNSTEHRSSSYHEN 833  
QY 760 ----- 759  
Db 834 DMENQSNITWYLLPLGPKSGNREODKPTIKTGRPHMKHRSWMKAPAGKTGRHGN 893  
Db 894 PKNYSYGMKSEEDIPSELLPKOKITSFLNRRMRYASEKSEYETIAANGEDTDVDKLTN 953  
QY 760 -----SFSONSRRPS-----TRO 772  
Db 954 SPONONITVPREGESISHNTTKKPSDLPFFSGVGHKSPHVRQDEBENSGOKROLFLRTKK 1013  
QY 773 KO----- 774  
Db 1014 KKKMKLALHSPLSPRGFDPLGNHNSPPDRRLNLHSLHLKSNETALSPDLNQTSPSM 1073  
QY 775 -----FNATPP 780  
5

Db 1074 STDRSLPDYNOYSKNDTEQMSSSLDYOSVPAEHSPTFPADDPQHTSTDDPSYRSSPP 1133  
QY 781 VLKR----- 784  
Db 1134 ELSQGLDYDLSHDFYDDTIGLISFFPPDQSQKSPSSDDQALPSSDLSLFTISPLDQTI 1193  
QY 785 ----- 784  
Db 1194 IYFDLQDLLSPEDNQKTSPPDLGVPLSPDDNQKTSPPDLGOVLSLSPDDNQKTSPPDLG 1253  
QY 785 -----HOR----- 787  
Db 1254 QVPLSLDDNQKTSPPDLGVPLSPDDNQMTSPDLGVPLSSDDNQKTSPPDLGVPLFPE 1313  
QY 788 -----EITRTTLOSDOEE----- 800  
Db 1314 DNQNTYLDLSQVPLSSDQNOETSSDLDLTLPDQOTYSPDLQPLPSDNSQVTVSPD 1373  
QY 801 ----- 800  
Db 1374 LSLTLSPDFNEIILAPDLGOVTLSPDLIQTNPALNHGKASADPDQASYPPDSQASS 1433  
QY 801 -----IDYDDT----- 806  
Db 1434 LPELNTLTPHDLTHIPPSPSPTLNNTSLSRKFNPLVYVGLSRVGDVETVPSSEPER 1493  
QY 807 ISVEKKEDFDIYDEDENQSPRSFQKTRH-----YFLAAYERLMDY 848  
Db 1494 IDEBIAEDDFYVYNDYRFDTDTDVNSSRNPTLAWTLRGHGKAKKRYTIAAEITWY 1553  
QY 849 GMSSEPHVLRNRAQSGVPO--FKVYVQEPETDGSFTQPLVGRBLNHLGLLPYIRAE 905  
Db 1554 AERQASEM--DHEBDGHPKOTTYKRYKRYLDSTFSRDBRAYEHLGILGPVIRAE 1611  
QY 906 VEDNIMVTFRNOASRPYSYSSLSYE-----EDRQQAEPKKNVKNENKITYRWK 957  
Db 1612 VDDVIOVRKKNLASRPYSLHAIGLSYKESSEKTYEDESPEWFODDAVQPNSSYTYWH 1671  
QY 958 VOHHNAPTDEPDKAMAYFSDVBLEKDVHSGLIGPLLVCHTNTLNPAGROVTVQEFAL 1017  
Db 1672 ATRKSRPEMPSGACRAMAYSAVNERDISHGLTLPILICRGTILAMERNLPMKREPLV 1731  
QY 1018 FTTIPEDEKSWYFTENEMERNCAPCIONEDPTEKENTRPHAINGYIMDTLPLGLMAODQ 1077  
Db 1732 LPMVFDEKSKSYEKS--KGSR-----RIESPEEKNAKFYALINQMIVN-LPGJRMVEOE 1783  
QY 1078 RIRWYLLMSGNENHSHIFESGHVFTYRKKEEYKALNVLPGVEYTEMPLSRAGIWRV 1137  
Db 1784 WVRHLNLNMGSSRDJHVHFGOTLLDNKTKOHLGVWPLRPGSFKYLEMKASRPGWMLL 1843  
QY 1138 ECLIGENLHAGSTLFLVYSNKCQTPPLGMAAGSHIDFQITASGOYQWAPKLARLHYSGS 1197  
Db 1844 DTEVGRNOYAGMOTPLLLIDKECKMPMGSLTGIVISDSQIKASEYITLYPERLARLNANNS 1903  
QY 1198 INAWSTKE-----PRS-WIKVDLAPMIIHGIKTQAGAROKESSLYISOFITMYSLDGCKW 1251  
Db 1904 YWAMSTEKALDLPKIPKWIQVDMQKEVVYTGIOGAKHYLKSCTPTEFOVAYSSDQNMW 1963  
QY 1252 QYTRGNSGTIAWFRGNVDSGIKHNIFNPRTIAYIRLHPHYISIRSTIKRELMLGCDLN 1311  
Db 1964 QIFRGKSGKSVYFTGNSGSGSTIKENRDLPRIVARYIRHPKYSYRNPRLTLELOGCEVN 2023  
QY 1312 SCSMPLGMEKSAISDAQUTASSYFTNMEAT--WSPSKARLHLQGRSNMARPVONNPKEWLO 1370  
Db 2024 GCSTPLGLELDGRIOKOTITASSFKKSMGDIYERPELARLNMAQGRVANAQANNNKOWLO 2083  
QY 1371 VDFQKTMKVYVTTQGVKSLTSMYVKEFLISSQDGHQWTLTFQNGKV--KVFQGNMDS 1428  
Db 2084 VDLKIKKVTALVTOGCKSLSEMENVKSYIOYSDQVAMKPYROKSSMYVCKFEGNSNT 2143  
QY 1429 FTVPVNLSLDRPLTRYLKRIHQSWYHQTALRMEVIGCAPODLY 1471  
Db 2144 KGHMKNFENPPLISRIRIIPKTMQOSIALRLELFGC---DIY 2183







```
Db 1697 LMSRPSLAHAGLSYKSSSGKTYEDDSDPEMFEDNAVQNSSTYVWMAHATERSGPPSG 1756
Qy 969 FQCKAAVFSVDVLEKDVHSLGLPLVCHTNTLNPAGHQVAVQEFALFETIDERSK 1028
Db 1757 SACRAMAYTSAVPEKDHSGLGLPLICKGTLKKNMPDREVLPLMVFDEKSS 1816
Qy 1029 YFTENNERNCRAPCNIOMEDPTEKREYRFAHNGYIMDLPLGLVMAODRIHMYLSMGS 1088
Db 1817 YXEKFTFSWR-----LTSSEVKNSHKFAHNGMIYN-LPGLMRYQEWVRLHLMLG 1869
Qy 1089 NENIHSIHSGHYFYTRKKEEYKALNLYPGVFYEYEMLPKAGITRVECLIGEHLHAG 1148
Db 1870 SBDIHVHHGQTLLENGTQOHQOLGWPLPLPGSFKTELEMTKSAGWMLDTEVEGENDRAG 1929
Qy 1149 MSTPLFVNSKQOTPLGMASGHIRDFOITASQYQGWAPKRLARLHSGSINAKTEKPPS 1208
Db 1930 KQTPFLIDRECKMPGLSLGLADSOIKASEFGWHQPLARLHNGGSIINAWIT-DKFS 1988
Qy 1209 -----WIKVDLLAPMIHIGITOGAROKFSSLSYISQITIMYSLDGKKQTYRGNSGT 1261
Db 1989 GESNSKPMIOWDMQREVFTGIOTGAKYLYKSYTTEFNVAVSSDQRMNRIFKGNSTKN 2048
Qy 1262 LWFVGNVSSSGIKNIENPPIARIIRLPHYSIRSTIRLMELMGCDLSCSPLGMS 1321
Db 2049 VWFVGNSDASTTENQOFDPVARYIRISPTESYNKPAKRLLEQGEVNGCSTPLGMS 2108
Qy 1322 KATSDAQITASSYFTNMFEAT-NGSPKARLHLOGRSNAMPQVNNPKMLVDFOKTKMYT 1380
Db 2109 GNIKNGQITASSFKKSKMGWDYEPFRARLNAQGRVNAWMAKANNMOMLOIDLKIKKIT 2168
Qy 1381 GVTTOGVKSLTSMYKFEKLSSSDGHOWTLFQNGV--KVFQGNDSFTTPVNSLDP 1438
Db 2169 ATTOGCKSLSEPMYKRYTIOYSDRGVEKSYREKSSMYDKIFEENNNIKGHYKNFNP 2228
Qy 1439 PLTRRLRIHPQSWHQIALRMVLCCEADLY 1471
Db 2229 PIIISREIRIIPKWNOSIALRLELFC--DIY 2258

RESULT 6
Q92024 PRELIMINARY; PRT: 1157 AA.
AC Q92024.
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE HEPHAESTIN.
GN HEPH.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RA MEDLINE=99140771; PubMed=9988272;
RA Vajpe C.D., Kuo Y.M., Murphy T.L., Cowley L., Askwith C., Libina N.,
RA Gipschier J., Anderson G.J.;
RT "Hephæstlin, a ceruloplasmin homologue implicated in intestinal iron
RT transport, is defective in the sla mouse.";
RL Nat. Genet. 21:195-199(1999).
DR EMBL; AF082567; AAD16035.1; -.
DR HSSP; P00450; 1KCM.
DR MGD; MGI:1332340; Heph.
DR InterPro; IPR001117; -.
DR InterPro; IPR002355; -.
DR Pfam; PF00394; Cu-oxidase; 3.
DR PROSITE; PS00079; MULTICOPPER_OXIDASE1; 3.
DR PROSITE; PS00080; MULTICOPPER_OXIDASE2; 1.
SQ SEQUENCE 1157 AA; 129681 MW; 411BD1DA2BFEC2FD CRC64;
```

```
Query Match 22.8%; Score 1792; DB 11; Length 1157;
Best Local Similarity 33.8%; Pred. No. 7e-130;
Matches 402; Conservative 171; Mismatches 418; Indels 200; Gaps 20;

Qy 20 ATRHYIGAVELMSDYQOSDLGELPDARPPRPVPSF-----PENTSVYKKTLEYEVFV 75
Db 24 AIRNYYLIGIDMOMNAPKGRNVIITQTLNNDVASSFLSKGRNIRGSSKRYKTYKEYSD 83
Qy 76 HLEFIAPRPMMGLIGPTQAEVYDPVTYTLKNMASHPSLAVGVSWKSEGEAYD 135
Db 84 GTYTEELAKPAMWGLFGLPLQAEVGYILLHLKNFASRPTTHPHGVFEKSEGSLYD 143
Qy 136 QTSOREKEDDKVPPGSHYVWQVLENGPMASDPLCLTYSYLSHVDLVKDLNSGLAL 195
Db 144 GSSGYLTKADDSVPPGSHVYVMSIPESHAPTEADPACLTIMYHSHVADARDIATGICPL 203
Qy 196 LVCREGLS-----AKEKQTLHKFLLFAVDEKSMHS-----TKNSLMDORAAASR 245
Db 204 ITCKRGTLDNSPPQRKDVDFHNEFLFSYIDENLSWHLNDNIATYCSDPASVDKEDGAFQ 263
Qy 246 AMPKMTVNGVYVNRSLPLGLICHRKSVYMHVIGMGTTPVHSHFLEGHFVLRNHRASL 305
Db 264 DSNRMAHINGFVGNLPELSMCAQKHVAHNLBGMGEIDVHTAFHFGQMLSTRGHTTVA 323
Qy 306 EISPTFLTAQTLMDLGOFLFCHTSSHQHDMEAIVKVDSCREBPOLRMKNNEABDY 365
Db 324 NIFPATFVTAEMVPQKSGTGLSCENVSHLRSQMAFFKYKCSMDPPV----- 372
Qy 366 DDOLITSEMVVAFDDDNSPSFIOISVAKKHKTWVHVAIEEEMWDAPLVLPDDR 425
Db 373 -DOLTG-----KVROYFLQAEHLOMDGP--IGYDGT 402
Qy 426 YKS-----OYLNGPQIRGKRYKRYEMAYTDEFTKTRALQHS--GILGPLYG 474
Db 403 GKSLREBGSPPDKYFOKSSSRIGTYMKRYAEFOBETQOERHODEEHLIDGLVIRA 462
Qy 475 EVGDTLLIFKNOASRPYNIYPHGTIDVRLPLYSRLPKGVK--LKDPILPGEIFRYK 531
Db 463 EVGDTIQVVEYNRASQPFSLQPHGVFEKKNSEGTAVNDGTSHPKVAKSF-----EKVY 517
Qy 532 WTVTVDEGPTKSPRCLTRYSFVMMERDLASGLGPLLICYSKESVDRGNQMSDRN 591
Db 518 WTVVPHAGPTADPACLTVMYFSAADPTDNTGSLGLPLVCAGLGDGKQKGVDEF 577
Qy 592 VILFSEVDENRSMYITENTIOFLPNPAG-----VOLEDE-FOASNIMASINGVYDFS 644
Db 578 FLIFTVFDEHESWYNNAN-----QAAGMDSLSLSEDEVGQDSNRMAINGLFESNLP 631
Qy 645 QLSVCLHEVAWYILSIGAOTFLSVFSGYTFKHKMYEDTLTLPFSGETVFMSENP 704
Db 632 RLDMCKGDTVAWMLLIGTETDVHGVFEGNTVOLGMRKGAVMLEPHTFVTAIQDPNP 691
Qy 705 GLWILLOCHNSDFNRBMVTLALKYSSCDKNTGDYEDSYDISAYLSKNAALERPSSFSQN 764
Db 692 GIFEIYCOAGSHREEMQALYVNSQ-----SSHQD----- 722
Qy 765 SRHPSTRQKOFNATPVLLKHKHOREITRTLQSDQDEIDYDITISVEMKKEDEIYDEDEN 824
Db 723 ----- 722
Qy 825 QSPRSFOKTRHYFIAVERLMDY-----GMSSPHYLRNRAOSGVSQPFK 870
Db 723 -SPRQHYQASRYVYIIMAEETEMDYCDRSWELEMHNTSEKDSYGVFLSNKGLGSKYK 781
Qy 871 KYVFOEFTQGSFTQPLRGEINHLDDLQPYTRAVEDIMVTFRNQASRPSTFS-SLI 929
Db 782 KAVFREYTDGTFRILPRRSGPEEHLIGLPLRGEVGDILLTVFKNKASRPSTIAHGV 841
Qy 930 SYBEDROGAEPKRNKVFKNETKTYFMKVOHNAHPKDFCDKAAVAFSDVDLEKDVSG 989
Db 842 ESNITGEPQAAE-----FGVLTITQWNIPEBSGEPSPSACVSVITYISANDPIKMG 894
Qy 990 LIGPLLVCHTNTLNPAGHQVTVQEFALFETIDETKSMYFTENNER-NCRAPCNIOMED 1048
```



Dd		895	IUGPLVINCNGLENGGNDMDREAFLLFLIFDEMGYSMLKENIATYTPQESHVNLMKD	954
Oy		1049	PFFKENEYRFAHNIYMDTLPGLVMAODORIRMYLWSGSNENIHFSGHVFYVRKE	1108
Dd		955	AFFLESNNKHAIINGKLIALRGITLVYQGERRAVMYMLAMQDDIDITHVHAESFYLQNQC	1014
Oy		1109	EKKMALYNLYPEVFETVEMLPSKAGIWRARECLIGELHNGMSTLFLYVSNK	1159
Dd		1015	STRADVLDLPETEFEVENYASNPGMTILHCHVTVDHVHAGMETIFTVLISHE	1065
<hr/>				
	RESULT	7		
	09J197			
ID	09J197	PRELIMINARY:	PRT:	1084 AA.
AC	09J197			
DT	01-OCT-2000	(TREMBLrel. 15, Created)		
DT	01-OCT-2000	(TREMBLrel. 15, Last sequence update)		
DT	01-MAR-2001	(TREMBLrel. 16, Last annotation update)		
DE	GPI-ANCHORED CERULOPLASMIN.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OX	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
NCBI_taxid=10116;				
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=SPRAGUE DAWLEY; TISSUE=BRAN;			
RA	MEDLINE-20127919; PubMed-10660599;			
RX	Patel B.N., Dunn R.J., David S.,			
RT	"Alternative RNA Splicing Generates a glycosylphosphatidylinositol-			
RT	anchored Form of Ceruloplasmin in Mammalian Brain."			
RL	J. Biol. Chem. 275:4305-4310(2000).			
EMBL:	AF202115; AAF34175.1; -			
DR	InterPro; IPR001117; -			
DR	InterPro; IPR002357; -			
DR	Pfam; PF00394; Cu-oxidase; 3.			
DR	PROSITE; PS00079; MULTICOPPER_OXIDASE1; 3.			
DR	PROSITE; PS00080; MULTICOPPER_OXIDASE2; 1.			
SQ	SEQUENCE 1084 AA; 123748 MW; 95D5B4154257C55 CRC64;			
<hr/>				
	Query Match	21.6%;	Score 1703;	DB 11; Length 1084;
	Best Local Similarity	32.2%;	Pred. No. 5e-123;	
	Matches 388;	Conservative 190;	Mismatches 427;	Indels 200; Gaps 23
Oy		5	LSTCFCLLRCCSFATRRIYGIAGVELSMDYMQ-SDLGEL-FYDAFPFRVVKSPFFNTS	62
Dd		6	LSALLF-HSSLAWTRERHYIGITEAWDVASGSEEEKLIVSDTEQSNFYLRNGPDRIG	64
Oy		63	VYYKTLTVEEFTVHLFNIAKPPPMGGLGPITQAEEYDVTVITLKNAASHPVSLAVGV	122
Dd		65	KRYKALISEYTDGTFYTKTIIDKPAPMLGLLGPIYIKAEVGDKSVHNKNFASRYTHAHVY	124
Oy		123	SYWKASEGAEYDDOTSQREKEDDKVYPGGSHTYVMQVLKENGPMASDPICLTYSYLSHVD	182
Dd		125	TYTKANEGAIYEDNTTDFQRADKDLFFPGQQYLYULRA-NEPSPEGGDSCNYTRIYHSVD	183
Oy		183	LYKDIINSGLIGLIVACRGSSLAKEKTOTL-KHFIILFAFDCKSMHSTKKSIM-----	236
Dd		184	APKDIASIIIGLIPLLICKKGSLHKKEENTDOEFVLFMSVYDENILSWYLEDNIKTCFCSEPE	243
Oy		237	-QDRDAASARAARKMHNTANGVYNRSLPGLIGCHKRSUYVHWVIGMTGTPEVHSIFLEGHTF	295
Dd		244	KYDKDNEDFOGESNRMYSLINGYFGSLPGLISMCAEDRRVKNKYLGRMGENDVDSHELPHGQL	303
Oy		296	LYRNHRQASLEISPTFFELTAQTLLMDLGOFLFLCHISSHQHDMGEARYKAVKVDSCPPEPOLR	355
Dd		304	TSKNYHTDIIINLFPATLLIDVSVYAONPGVWMSCONTNLHKLKGAIFAFFOVRRC-----	356
Oy		356	MKNNEAEYDDDLDESMDVVRFPDDNSPSTIQIRSVAKKPKRKPMVHHVIAAEEDMDVA	415
Dd		357	---AKPSP--DDDIDDHNV-----RH-----YIAAEETIMDWA	385

OY	146	P-----LVLADDDSSYSQYLNNROKRLGRKKYKVRMAATDCEP---KTRALI	461
OY	416	-----LVLADDDSSYSQYLNNROKRLGRKKYKVRMAATDCEP---KTRALI	461
Db	386	PSGDTFTGEMNLISGSDSRVFEQ---GATRIIGSYKKLVREYTDSTFNKORCPD	441
OY	462	QHESGILGRLPYGEGDTLLILFEKNQASRPYNTIYPHGITDVR----PLYSRRLPKGVKHL	517
Db	442	EHNLGILGRLPYMAVGOLINTYFNKKQPRLSLQPMGVRTTKNECTYTG---PDRSSK	498
OY	518	KDFPLPGLGELFKKYAVTVEDGPTKSDPRCLTRYSSPVMMERDLASLIGRLLCYKES	577
Db	499	QASHVAPREFTETTYEMTVPEKKEGPTPYADPCLSKMYSGVDLTGDTFTGLGIPKICKGS	558
OY	578	VDQNGNQMMSKRNVLIFSVEDEKRSYUYLENIQRLPNAGSQLDEDPFOASNTMHSN	637
Db	559	LLADGQRKDXDKKEEYLFATVFEDESNELLDLDNIMFTTAPEVNDKDEDFQESKNKHSN	618
OY	638	GYVDSLSQ-LSVCLHEVAWYWLISGAQDPLSVFSGYEFKKHMYEEDTLTFPESGET	696
Db	619	GFMIGNLPGLMCMGIBETIWTYLFSGAGADVHGIITSGNTIYLSKGRRTANLFRPKSLT	678
OY	697	VFMSENMGLWILICHSNDFERNKMTALLKVSSCDKNITGDYEDSEDIISAVLLSKNAI	756
Db	679	LLMTPTFEGSPDVECLTTHDHYTGMMOKYUYNOC-----KGQEDDTLV-----	722
OY	757	EPRSFQNSRHPSTROKQFNATPVLKRNQREITRTTLOSQDEIDYDITISYEKKEDF	816
Db	723	-----	722
OY	817	DIYDEDENQSPRSFOKKTNYFTAAVERLMDYGMSSSPHYLRNRAQSGV-----	866
Db	723	-----QCERTYUUAAYEVENDYSPSRDMEKELNHLIDQVNSNAFLDKEEF	767
OY	867	--POFKKVVFOEFTDTSFTOPRLRGELENLILGRTYRAEVEDNIWTFRRNQASRPS	923
Db	768	FIGSKYKKVYRRETTDSTFRFQYKRRAREENHLILDRLLHADGVKKYKAFFNMASRPS	827
OY	924	FYSSLSISTEEDQROGAERKKNFYK--PNETKYEFKVOHNMAPTKDEDFCKAMAFSDV	980
Db	828	IHA-----HGVTKSSSTVPTPLPGEVRTYIQDIPERSGAGTEDSPCIIPMAYSTV	877
OY	981	DLEMDVHSGILGRLPYC---HTNTLNPANHRQUTVQERPLFTIPRETKSWTFENMEEN	1037
Db	878	DRVADLTSGILGRLPYCRKSYKVVKNPK-----KKMEYSILFLVPEBNSWYLDONINTY	932
OY	1038	CRACNLQMEDPREFKENYRPHAINGYIMDTPLGLVAODQRIKLYLLSMGSNENHSIHF	1097
Db	933	SDHREKYNKQDEEFTIESKNKHAINGKRFGLQSLTHYHVDDEVWYVMAMANGNELDLHTYHF	992
OY	1098	SGHVFYRKKEEKYKMALYNLYPGVFETVEMLPKSCAIGWRVECLIGENHLHAGNSTFLYVS	1157
Db	993	HGHSEFOYKRNHRGHSIDVEFDLFCGYOTLTFMFROTPTGTMLLHCHVTGHDHIIHAGVTTYULP	1052
OY	1158	NKCOF 1162	
Db	1053	NQASS 1057	
RESULT	8		
O9XT27			
ID	O9XT27	PRELIMINARY;	PRT; 1048 AA.
AC	O9XT27;		
DT	01-NOV-1999 (Tremblarel, 12, Created)		
DT	01-NOV-1999 (Tremblarel, 12, Last sequence update)		
DT	01-OCT-2000 (Tremblarel, 15, Last annotation update)		
DE	CENUOPLASMIN.		
OS	Ovis aries (Sheep).		
OC	Eumariyota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;		
OC	Bovidae; Caprinae; Ovis.		
OX	NCBI_TaxID=9940;		
FN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=LIVER.		



RA Lockhart P.J., Mercer J.F.B.,  
 RT "Cloning and Expression Analysis of the Sheep Geruloplasmin cDNA,"  
 RL Submitted (Mar-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF134814; AAD41477.1; -  
 DR HSSP; P00450; 1KCM.  
 DR InterPro: IPR001117; -  
 DR InterPro: IPR002355; -  
 DR Pfam; PF00394; Cu-oxidase; 3.  
 DR PROSITE; PS00079; MULTICOPPER\_OXIDASE1; 3.  
 DR PROSITE; PS00080; MULTICOPPER\_OXIDASE2; 1.  
 SQ SEQUENCE 1048 AA; 119125 MW; 925f16d7b0549cbb CRC64;

Query Match 20.6%; Score 1619.5; DB 6; Length 1048;  
 Best Local Similarity 31.8%; Pred. No. 1.4e-116;  
 Matches 383; Conservative 180; Mismatches 422; Indels 221; Gaps 27;

QY 9 PFLC-LRRC-----FSATRRYLGNVELSMVQNSDLGE---LPVDAFRPRPKSPFN 60  
 DB 4 FLCLFLGLGTSVWAKDKHYYIGILETAMNY-ASDHAEEKLISVTEHSNIYLQNGPNR 62  
 QY 61 TSVYKKTLEFEFTVLEFNIAKPRPMGLGPTIOAEVYDVVITLKNMASHPVSLHAY 120  
 DB 63 GSVYKKAAYLOYTDENFTVIEKRPVWLGFLGPIIKAEIGDKVYVHLKNFASRPYTFNAH 122  
 QY 121 GVSYKASEGAEYDDQTSOREKEDDKVPFGSGHTYVYQVLEKNGPMASDPLCLTYSYLH 180  
 DB 123 GLTYKKEHGAAYPDNTDLOKADKDVQGEQCLYILHANPEQGEEDENCYTRIYHSH 182  
 QY 181 VDLVMDLNGSLGALLVCEGSLAKEKKTQTLHK-FILLFVDPGKSMH-----SE 230  
 DB 183 IDAPKDIASGLPLGRLHCKKDSLDEBEKKNIDKEFYVMSVDENTSWILEENIKTYCSE 242  
 QY 231 TKNSLMODRDA SARAPMKHTVNGYVNSRLPGLICHRKSVYWHYIGMTPTPEVHSIFL 290  
 DB 243 PEKVQDNEDFQESN---RMYSVNGYAFGLSLPGLSMCAEDRKYVWLEFGMNEIDVHAHF 299  
 QY 291 ECHTFLVRHHRQASLEISITPFLTAOTLLMDLGLFCHISHQHDGMAYKVDSCPE 350  
 DB 300 HGQVLTSKNYRVDTINLFPAFLFEDAFVAVQNPQWMLSCONLHLKAAGLOAFWVQDCKK 359  
 QY 351 EPQLMKNEAEADYDDDLTDEMDVVRFPDDNSPSFIQIRSAVKHHPKTVIYAAEE 410  
 DB 360 SS-----SEDN-----IHGKNVNH-----YIAAEV 361  
 QY 411 DWDYAPL-----VLAPDDRSYKSOYLNNGPORIGRKYKKVRPMAYTDETF---KTR 458  
 DB 382 IWNYPGSDIAFTKELRAFGSAS--EAFEDGPTRIGSYKKLYVREYTDASFQNGKER 439  
 QY 459 FAIQHESGLGPLYGEVGDLLITFKNOASRPYNIYPHCI-----TDVAPLYSRL 510  
 DB 440 GPEEHLLGLGPIVIAAEVGDITRVTFHNKAAPLSTIEPIGVARDKNNEGTYSPTSGCP 499  
 QY 511 PKGVHKLDPFLPGEIFYKMTVTVEDGPTSDPRCLRYYSFVNMEDRLASGLIGPL 570  
 DB 500 PSG-SH-----VAPKGTFTYEMTVKEVGPPTYKDPVCLAKMYIS--GSTRKIDITGLIGPM 551  
 QY 571 LICYKESVDQNGQIMSDKRNVLFSVPDENSNWLTENIQRLFPNPAVOLDEDFEFOAS 630  
 DB 552 KICRNGSLANRLKNVDEFLFTVPEDENSLLDDNINKMTTAPDQVCKNEDEFOES 611  
 QY 631 NIMHSINGVFPDLO-LSYCLHEVAYWYTLSTGAQTDPLSVFSSGYTFPKHKNVYEDTLTL 669  
 DB 612 NKMHMNGMYNONGPLSLSCQDSDVMWYLFASGNEVDIGHIYFSGMTYISRGERRDTANL 671  
 QY 690 PPFSEGEYFMSMENPGMLTGLGCHNSDFRNKGTALALKVSSCOKNTDDYEDSDISAYL 749  
 DB 672 FPQIQLSLFPMQDITAGTFVEYECITTDHYTGAKOKATYSQCGORS-----EDLYLTL 723  
 QY 750 LSKNNAIEPRFSFONSRRHSTROKOPNATPPVLKRHOREITRTTLOSDEIDYDPTISV 809  
 DB 724 ----- 723

QY 810 EMKKEDFDIYDEDEENOSPSPFOKTRHRYFIAVERIMDYGMS-----SSPHVLNRNAQSGS 865  
 DB 724 -----GERTYYIAAEVEMDPSPSKWEKELHNLQEOULSVA 760  
 QY 866 V-----POKRVYFOEFTDGSFTOPLYRGELNENHGLIGLPYRAEVEDENIMWTFRN 916  
 DB 761 FLDKKEEYIGSKTKKYYVRFDTSTFOVYERKEEGEENHGLIGLPOLHADVGKVNITRN 820  
 QY 917 QASRPYSFYSLSLISYEDOROGAEPKRNKPVKNETKYTFWKVQHNNAPPKDEFDCANAY 976  
 DB 821 MATRPYSIHMHGKTESST---VTP-----TAPGETRTYIWKIPERSGAGMGSDSPICPMVY 873  
 QY 977 FSDVLEKDVHSLGIPRLVCHTNTL---NPARGOVYQOEALPFTTIDFKSYTFEN 1033  
 DB 874 YSTVDYRKDFSLGIPLYCRKHNYLVKSNP-----IKRLFSFLFLVFEDENSWYLDN 928  
 QY 1034 MEKNCRAPCNIQMEDPTFKENYRFHAINCYIMDTLGLVMAQDQRIRWYLLSMGSNENH 1093  
 DB 929 IKTYSDHPKRVKANEEMESKMHAINGRMGNLGLTM-----HVGNEVDLH 977  
 QY 1094 SIHPSGHVFTVRKKEEKYKALYNLYPGVFEYEMLPSKAGIWRVCLIGELHLAGMSTLF 1153  
 DB 978 SVAFHGHSFQYQHRGITYSDVFDLPFGTYQTLEMPKTPGIWLLCHYVDHIIHAGMETTY 1037  
 QY 1154 LVYSNK 1159  
 DB 1038 TVLPNE 1043

RESULT 9  
 ID 075180 PRELIMINARY; PRT; 891 AA.  
 AC 075180;  
 DT 01-NOV-1998 (TREMBLrel, 08, Created)  
 DT 01-NOV-1998 (TREMBLrel, 08, Last sequence update)  
 DT 01-MAY-2000 (TREMBLrel, 13, Last annotation update)  
 DE KIA0698 PROTEIN.  
 GN KIA0698.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=BRAIN;  
 RX MEDLINE=98403880; PubMed=9734811;  
 RA Ishikawa K., Nagase T., Suyama M., Miyajima N., Tanaka A., Kotani H.,  
 RA Nomura N., Ohara O.;  
 RT "Prediction of the coding sequences of unidentified human genes. X.  
 RT The complete sequences of 100 new cDNA clones from brain which can  
 RT code for large proteins in vitro.";  
 RL DNA Res. 5:169-176(1998).  
 DR EMBL; AB014598; BAA31673.1; -  
 DR HSSP; P00450; 1KCM.  
 DR InterPro: IPR001117; -  
 DR InterPro: IPR002355; -  
 DR Pfam; PF00394; Cu-oxidase; 3.  
 DR PROSITE; PS00079; MULTICOPPER\_OXIDASE1; 3.  
 DR PROSITE; PS00080; MULTICOPPER\_OXIDASE2; 1.  
 SQ SEQUENCE 891 AA; 100624 MW; AB70D0436109C5B2 CRC64;

Query Match 17.0%; Score 1334; DB 4; Length 891;  
 Best Local Similarity 30.7%; Pred. No. 1.6e-94;  
 Matches 321; Conservative 148; Mismatches 358; Indels 220; Gaps 23;

QY 250 MHTVNGYVNSRLPGLGCHRSKYVYWHYIGMTTPYVHNSIFLEGHFLVARNHQAASLEIS 309  
 DB 1 MAININGVFNGNPELNMCAQRYVAHMLFGMGNEIDVHTAFHFGQMIITRGHTDVANLFP 60  
 QY 310 IFFLTAAOTLLMDLGOFLFLCHITSHOHGMEAYVVKVDSCEPPEPOLRMKNNEAEYDDDL 369  
 DB 61 AFTVTAEMVPEWEGTWLISQVNSHFRDGMQALYKVKSCSNAPV-----DLL 108



QY	370	TDSEMDVVRFDSDNSPFIQIDRSAAKHNRPTWHTXIIAAEBEDMDYAPL-----VLA	420
Db	109	TG-----TG-----KVQFTEAHBIDIOVDYPMCHDSTGKNKE	141
QY	421	PDDRSRYSOYLNNNPORIGIRKRYKVRPMATYDEFTKTRERAIQHS--GILGPLYGVEVD	478
Db	142	PG--SIDKFFQKSSSSIGSTYWKVREARQODETPQEKHMLBEDRHLGILGYIAREVD	199
QY	479	TLLIIFKQASRPYNIYPHGITDVRPLYSRRLPKYVKNLKDPEILGELFK-----XKW	532
Db	200	TIQVFFYNRASQPSRMPQHV-----FYEKDYEGTV--YNDGSSYGLVAKDFEKTATYRW	252
QY	533	TVTVEDDPTSDRPRCLIRYSSFSVNMNERDLASGLIPRLILCYKESVDQNGNIDMKRV	592
Db	253	TVPRHAPETADPRCLITWMTFSAADPLRDTNSLVGPRLVLCRAGAGADKQKGVDFEEF	312
QY	593	ILFSEVPENSNMYTENTIQFLEPNPAGVOL-----EDPE-FQASNIMHSINGVYFSL-	644
Db	313	LLEFVLDENCSWYSN-----ANQAAAMLDPELLSEDLIEGQDSDRHMIAINGFLESNP	365
QY	645	QLSVCLHEVAWYIILSIGAOTDPLSVFFSGYTERKHKMYEDTLTLEPFSGETVMSMENP	704
Db	366	RIDMKCKDPTAAMHNLIGTETDVGWVFCQNTVOLGMRKKAAMLEPHTFVMAIQPDUL	425
QY	705	GLMILTGHNDSFRNRGWTALLKVSQCDKNTGDIYEDSYEDIASYLLSKNNALERPSPSON	764
Db	426	GTEFITYQAGSHRAGRAIYNSQC-----	451
QY	765	SRHSPTKOKOPNATPRPLKHNOREITRTTLOSDEEIDYDPTISVBMKKEDEYDEDEN	824
Db	452	-----PQHQ-----	455
QY	825	QSPRSFOKTRHFIFAAVERLMDYGSSS-PHYLRNRAQSGV-----POFK	870
Db	456	ATPRQRQAARIYYIIMAEVEWYDCPDRSEREMHNOSEKDSYGLFSLNKGDLGJSRYK	515
QY	871	KVVFQEOETDSFTQPLRGSLNHLGILGYIAREVDNIMVTFRMOASRPSYFS-SLI	929
Db	516	KAVERETDGTFRILPRPRTGREELHGLGLPLKGEVGDILTYVFKNNASRPYSVHAHGVL	575
QY	930	SYEEDROGAAPRKNEFKPMETKTYFWKVQOHNHAPTKDEFDCKAMAYFSDVLEKDVHSG	989
Db	576	ESTTVWPLAAE-----PEVVTYQOMNIPERGSRPNDNSACVSWIYSAVDPIIDMSG	628
QY	990	LIGLVLCHNTLTPAHGROYTVQERALFETTPDETCKSWYFTENM-ERNCRAPCNIOMED	1046
Db	629	LVGSLAICQGIIEPHGGRSDMREAFLLFLIDENKSWYLEBNVATHGSODPGSINIOD	688
QY	1049	PTFEENRPHANINGYIMDTPLGVLMAODORIRYLLTSLMSNENHISIHFSGHVFIYRKKE	1100
Db	689	ETFLIESKMHAIINKLANTLRGLTMOGGEVAVMTMLAMQODVLIHTHNAESFLYRNGE	748
QY	1109	EYKMAIYNLYPGVEFTEVEMLPSKAGIMRWECCLIGEHLHAGMSTLFLVYNSNKOTPLGNAS	1166
Db	749	NYRADVYDLDPRGFEVVEVMAVNSPGWIMLHCHVTDHVHAGMETLFLVEFR-----T	799
QY	1169	GHIDPOITASGOYGOMAPKALALHSGSINAMSTEEPRPSWIKVDLLAPMIIHGKTQGA	1222
Db	800	EHLSPPLVITKETEKAPRPDIE-----EGNVKMLGMIQIPK--NVEMLASVLV-----	846
QY	1229	RQKFSLYISOFIIMYSLDCKMKQOTR	1255
Db	847	-----AISVILLVLVLAAGSVWYQHR	868
RESULT	10		
ID	075659		
AC	075659:	PRELIMINARY:	PRT: 782 AA.
DT	01-NOV-1998 (TREMblrel. 08, Created)		
DT	01-NOV-1998 (TREMblrel. 08, last sequence update)		
DT	01-MAY-2000 (TREMblrel. 13, last annotation update)		

DE DJ466181 (COAGULATION FACTOR V (ACTIVATED PROTEIN C COFACTOR),  
DE COAGULATION FACTOR VIII (PROCOAGULANT COMPONENT) AND CERULOPLASMIN  
DE (EC 1.16.3.1, FERROXIDASE) LIKE) (FRAGMENT) .  
GN DJ46618.1.  
OS Homo sapiens (Human) .  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Wilkinson J.;  
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AL030998; CAA19742.1; - .  
DR HSSP; P00450; 1KCW.  
DR InterPro; IPR001117; - .  
DR Pfam; PF00394; Cu-oxidase; 2.  
DR PROSITE; PS00079; MULTICOPPER\_OXIDASE1; UNKNOWN\_2.  
FT NON\_TER 1 782 1  
FT NON\_TER 782 782  
SQ SEQUENCE 782 AA; 87939 MW; 2FBD2824CC19D2AB CRC64;

Query Match	15.4%;	Score 1214;	DB 4;	Length 782;
Best Local Similarity	32.2%;	Pred. No. 2.6e-85;		
Matches 292;	Conservative 121;	Mismatches 296;	Indels 198;	Gaps 18;

0Y	65	YKTLFVFEVTHLEIAKPRRPMNGISPTTQAEVUTQYTLKNNASHRYSIAVSV	124
Db	13	YKTLTYKKEIKDDSTIDEVAORAMGLFGRVLOAEGBVILHNLNFKRPTTIPHCYV	72
0Y	125	WKASGAEYDQTSOREKEDKVRPGGSHYVMVLKENGFMASDRPLCTLYSYSHDLY	184
Db	73	EKDSGSLYPRGSSGPLKADDSVPRGSGHIYNMTLIPGHAFTPADPACLTIMYISHDAP	132
0Y	185	KDLNSGLIGALLVCREGSLAEKQTLKFLFLFAVDEGKSNH-----SETKNSLMOD	238
Db	133	RDIAGLGLIPLCTCKRG-----THDFFLEFSVYDEMLSHMLNENIATYCSDPASVD	184
0Y	239	RDASARAMPKMHVGVNYSRLGLJGCHKRSUYMHVIGMTPREHSTFLBSGHTFLVR	298
Db	185	KEDTFQSSNMHAINGVREGNLRBELNMCQAKRRAMHLFGMGNIDVHTAFPHQMLTTR	244
0Y	239	NHRQASLEISPTTFLTAOTLLMDJGOLFLECHISSHOHDGNEAVYKVDSCREPERMKN	358
Db	245	GHNHDVANIFPATFETJAMVWMEGWTLSICOVNSHFRDGOALYKVKSCSMAPV----	300
0Y	359	NBEADYDODLIDSEMDVRRDDDNSPFIQIRSVAKKHRTWHTYLAAREEDMDVAPL-	417
Db	301	-----DLITG-----KYQYTFEANEIOMDYPRMG	325
0Y	418	-----VLAPDDRYSYQYLYNNGPORIGRKRYKVRFMATDETFTTREATIOHES--GI	467
Db	326	HDSGTGKMLRPRG--SISDKFFOKSSSRIGTGYKVRYEAFODETPQEKMLIEDRIIGI	383
0Y	468	LGRLLYGEVGTLLILFKNQASRKYNIYPHGITDVRPLYSKRLRGVYKHLKDFPLRGEI	527
Db	384	LGPVIRAEVGTIOVYENRASOPFSMOPHGV-----FYEDKEGTV--YNDGSSYGLV	438
0Y	528	FK-----YKMTVYVEDPRTSPDRCLTRYFSSFVNNEBRDASLILPRLICYKEVDOR	581
Db	437	AKPFEKYIYRMYTPVRPHAGPTADQACLTMMYFFSAADRIROTNSLTVARPLIVCRAGLAD	496
0Y	562	GNOIMSKRWVILFESVDENRSWLTENIQORFLNPAVOL-----EDPE-FOASINIH	634
Db	497	GKQGVXDPEFLFLFTVDENKSWSN-----ANOQAAMIDFLRLSEDLIEGODSNRMH	549
0Y	635	SINGVPRSL-OLASJCHEVAYVYILSIGAOTDLFVFSGYTRPKHMAVEDTLTPRFS	693
Db	550	AINGFLPSNLRPLDMCKGDYVAMHLLIGTETDVHGVMFOGQNTYQLOGMKRGAAMFLPHT	609
0Y	694	GETVFMSENENGLMILGCHNSDFRNRGNTALLKVSQCDKNTGTYEDSYEDIAVYLLSKN	753
Db	610	FVMAIMODDNLGTEIFYCOAGSHHEAGKRAIYVNSQC-----G466	



```
QY 754 NAIEPRSESONSRHPSHSTROKQENATPPVLKRHOREITRTTLOSDOEIDYDITISVEMKK 813
D 647 -----PGHQ-----
QY 814 EDFDIYDEDENOSPSPFOKKTTHRYFIAAVERLMDYGMSSS-PIVLLNRAOGSGV----- 866
D 651 -----ATPRORQOARIIYIMAEVEEDWYCPDRSWEREMWNOSEKDSYGFILS 699
QY 867 -----POFKVVFQEFIDGFTOPLYRGELNHLGLGPYIRAEVEDINWTFRRQAS 919
D 700 NKDGLLSGKKAAYKREYIDGTFRIIPRPTGPEHLGIILKGEVGDILIVYFKNAS 759
QY 920 RPYSFYS 926
D 760 RPYSVHA 766

RESULT 11
Q14286 PRELIMINARY; PRT; 216 AA.
ID 014286:
AC 014286:
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE COAGULATION FACTOR VIII ASSOCIATED PROTEIN B.
GN F8B.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93052386; PubMed=1427887;
RA Levinson B., Kendrick S., Gamel P., Fisher K., Gitschier J.;
RT "Evidence for a third transcript from the human factor VIII gene.";
RL Genomics 14:585-589(1992).
DR EMBL; M90707; AAC58466.1; -.
DR HSSP; P00451; ICFG.
DR InterPro; IPR000421; -.
DR Pfam; PF00754; F5_F8_type_C; 1.
DR PROSITE; PS01285; FA58C_1; UNKNOWN_1.
DR PROSITE; PS01286; FA58C_2; 2.
DR SMART; SM00231; FA58C; 1.
SQ SEQUENCE 216 AA; 24641 MW; 6C82D4F89E35A376 CRC64;

Query Match 14.0%; Score 1104; DB 4; Length 216;
Best Local Similarity 100.0%; Pred. No. 1,le-77;
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1264 VFFGVNDSSGKIHNFNPPIIARIYIRLPHTHYSIRSTLRMELMGCDLNSCSMPLGMSKA 1323
D 9 VFFGVNDSSGKIHNFNPPIIARIYIRLPHTHYSIRSTLRMELMGCDLNSCSMPLGMSKA 68
QY 1324 ISDAQITASSYFTNMFAITWSPSKARLHLGSRNAMPQVNNPKEMQLQVDFOKTKMKTGYT 1383
D 69 ISDAQITASSYFTNMFAITWSPSKARLHLGSRNAMPQVNNPKEMQLQVDFOKTKMKTGYT 128
QY 1384 TOGVKSLTSMYVKEFLISSDQGHQWTLFFONGKVKYVQGNDSFTPVVNSLDPPLLTR 1443
D 129 TOGVKSLTSMYVKEFLISSDQGHQWTLFFONGKVKYVQGNDSFTPVVNSLDPPLLTR 188
QY 1444 YLRHPOSVMHQAIRMEVLGCEADLY 1471
D 189 YLRHPOSVMHQAIRMEVLGCEADLY 216

RESULT 12
Q09RLX9 PRELIMINARY; PRT; 463 AA.
AC Q09RLX9:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
```

```
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE MILK FAT GLOBULE GLYCOPROTEIN MFG-E8 LONG FORM.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BAH/C; TISSUE=MAMMARY GLAND;
RX MEDLINE=99120894; PubMed=9920772;
RA Oshima K., Aoki N., Negi M., Kishi M., Kitajima K., Matsuda T.;
RT "Lactation-dependent expression of an mRNA splice variant with an exon
RT for a multiply O-glycosylated domain of mouse milk fat globule
RT glycoprotein MFG-E8."
RL Biochem. Biophys. Res. Commun. 254:522-528(1999).
DR EMBL; AB021130; BAA35180.1; -.
DR HSSP; P00740; IIXA.
DR InterPro; IPR000421; -.
DR InterPro; IPR000561; -.
DR InterPro; IPR001092; -.
DR InterPro; IPR001438; -.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00754; F5_F8_type_C; 2.
DR PRINTS; PR00010; EGFBL00D.
DR PROSITE; PS00022; EGF_1; UNKNOWN_2.
DR PROSITE; PS01186; EGF_2; 2.
DR PROSITE; PS01285; FA58C_1; 2.
DR PROSITE; PS01286; FA58C_2; 2.
DR SMART; SM00181; EGF; 1.
KW EGF-like domain; Glycoprotein.
SQ SEQUENCE 463 AA; 51269 MW; D719D2BE090E6427 CRC64;

Query Match 8.4%; Score 664; DB 11; Length 463;
Best Local Similarity 42.5%; Pred. No. 5,le-43;
Matches 136; Conservative 62; Mismatches 110; Indels 12; Gaps 5;

QY 1157 SNKCOTPLMASGHIRDFOITASGOY-----GONAPKLARIHYSGSINAW--STKEPFS 1208
D 145 ASKCSOTPLMEGALINDSISASVYMGLOIRNGPELARILYKRGYINAWATWSYDCKP 204
QY 1209 WIKVDLLAPMIHIGIKTGAROKFSSLYISQFIIMSLDGKKWQTYRGNSGTGLWVFQGN 1268
D 205 WIQVNLIRKMRVSGVWTQASRASRAEYLTKEKVAVSLDGRKFEFIODESGD-KEFLGN 263
QY 1269 VDSGSGIKHNFNPPIIARIYIRLPHTHYSIRSTLRMELMGCDLNSCSMPLGMSKAISDAQ 1328
D 264 LDNNSLTKVMKPEFTEAOYIRLVPVCHRGCTLRRELLGCELHGSGEPLGKNNTPDSQ 323
QY 1329 ITASSYFT--NMFA-TWSPSKARLHLGSRNAMPQVNNPKEMQLQVDFOKTKMKTGYTQ 1385
D 324 MSASSSYKTWNLRAPGWYHLGRLDNGKINAWTQASNAKEMQLQVDTGQRQVGIITQ 383
QY 1386 GYKSLTSMYVKEFLISSDQGHQWTLFFONGKVKYVQGNDSFTPVVNSLDPPLLTRYL 1445
D 384 GARDFGHIQYVASYKVAHSDDGQWTVYEQGSKVFQGLDNNSHKKNIFEKPEMARVY 443
QY 1446 RHPQSWMHQAIRMEVLG 1465
D 444 RVLPSVMHNRITRLLELGC 463

RESULT 13
Q043854 PRELIMINARY; PRT; 480 AA.
AC Q043854:
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE INTEGRIN-BINDING PROTEIN DELI PRECURSOR.
GN DELI.
OS Homo sapiens (Human).
```







```
DB 206 DGRKEFEIDDESGD-KEFLGNLDNNSLKVMNPFLEAOYLRLPYVSCRGCTLRPELL 264
OY 1307 GCDLNSCMLPMEKSKAIDAOITASSYFT--NMFA-TNPSPSARHLLOGSNAMRQVN 1363
DB 265 GCELHGCSPEPLGLKNNTITDSSOMASSSYKTWNLRFGWYPHLGRDNOCKTIANWAOQN 324
OY 1364 NPEKMLQVDFOKTMKVYVTTQGVKSILTSMYVKEFLISSODGHQWTLFFONGKVKVFO 1423
DB 325 SAKEMLQVDLQYGRQVYTGITGARDFGHQLQYVASYKVAHSDQGVQWTVYEEOGSSKVFQ 384
OY 1424 GNDSEFTPVNSLDPPLRLRYLRHPQSWHQIALRMEVLGC 1465
DB 385 GNLDNNSHKNIEKPFMARVYVLPVSWHNRITLRLELGC 426

RESULT 15
ID 035474 PRELIMINARY; PRT; 480 AA.
AC 035474; 035475;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, last annotation update)
DE INTEGRIN-BINDING PROTEIN DELT PRECURSOR.
GN EDL3 OR DEL1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND CHARACTERIZATION.
RC TISSUE=EMBRYO.
RA MEDLINE=98083109; PubMed=9420328;
RA Hital C., Zupancic T., Penta K., Mikhail A., Kawana M.,
RA Quettermous E.E., Aoka Y., Fukagawa M., Matsui Y., Platika D.,
RA Auerbach R., Hogan B.L.M., Snodgrass R., Quettermous T.;
RT "Cloning and characterization of developmental endothelial locus-1: an
RT embryonic endothelial cell protein that binds the alphabeta3 integrin
RT receptor.";
RL Genes Dev. 12:21-33(1998).
CC -1- FUNCTION: PROMOTES ADHESION OF ENDOTHELIAL CELLS THROUGH
CC INTERACTION WITH THE ALPHA-V BETA INTEGRIN RECEPTOR. INHIBITS
CC REGULATION OF VASCULAR-LIKE STRUCTURES. MAY BE INVOLVED IN
CC DEVELOPMENT OF VASCULAR MORPHOGENESIS OR REMODELING IN EMBRYONIC
CC DEVELOPMENT.
CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR MATRIX.
CC -1- ALTERNATIVE PRODUCTS: TWO ISOFORMS: LONG ISOFORM (SHOWN HERE) AND
CC SHORT ISOFORM, ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN ANGIOBLASTS AND EARLY ENDOTHELIAL
CC CELLS. BY EMBRYONIC DAY 13.5, ALSO EXPRESSED IN A RESTRICTED GROUP
CC OF NON-ENDOTHELIAL CELLS INCLUDING CHONDROCYTES AND RETINAL
CC NEURONS.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN THE EMBRYO FROM DAY 7. AFTER DAY
CC 15.5, EXPRESSION DECREASES AND DISAPPEARS COMPLETELY BY THE TIME
CC OF BIRTH.
CC -1- SIMILARITY: CONTAINS 3 EGF-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.
CC EMBL: AF031524; AAB86585.1; -.
CC EMBL: AF031525; AAB86586.1; -.
CC HSSP: P00740; 1EDM.
DB MGD; MGI:1329025; Ed113.
DR InterPro: IPR000152; -.
DR InterPro: IPR000421; -.
DR InterPro: IPR000561; -.
DR InterPro: IPR000742; -.
DR InterPro: IPR001881; -.
DR Pfam: PF00008; EGF_3.
DR Pfam: PF00754; F5_F8_type_C_2.
DR PROSITE: PS00010; ASX_HYDROXYL; UNKNOWN_1.
DR PROSITE: PS00022; EGF_1; UNKNOWN_3.
DR PROSITE: PS01186; EGF_2; 2.
DR PROSITE: PS01187; EGF_CA_1.
DR PROSITE: PS01285; FA58C_1; 2.
DR PROSITE: PS01285; FA58C_1; 2.
```

```
DR PROSITE: PS01286; FA58C_2; 2.
DR SMART: SM00231; FA58C; 1.
KW EGF-like domain; Alternative splicing; signal; Developmental protein;
KW cell adhesion; Repeat; Vascularization.
FT SIGNAL 1 16
FT CHAIN 17 480
FT DOMAIN 26 59
FT DOMAIN 78 116
FT DOMAIN 123 154
FT DOMAIN 161 311
FT DOMAIN 322 473
FT SITE 96 98
FT DISULFID 26 37
FT DISULFID 31 48
FT DISULFID 50 59
FT DISULFID 78 89
FT DISULFID 83 105
FT DISULFID 107 116
FT DISULFID 158 314
FT DISULFID 301 305
FT DISULFID 319 476
FT VARSPIC 218 221
FT VARSPIC 222 480
SQ SEQUENCE 480 AA; 53740 MW; 4CD91EEF261714D CRC64;
```

Query Match 8.3%; Score 655.5; DB 11; Length 480;  
Best local Similarity 44.6%; Pred. No. 2,5e-42;  
Matches 145; Conservative 45; Mismatches 120; Indels 15; Gaps 5;

```
OY 1159 KCOTPLGMSGHRDPOITRASGOY-----GQAPPLALRLHSGSINAMSTE--PFSWI 1210
DB 157 KCSGPIGIEGIIISNOOITASSSTHRLFGLRKYPPYARLNKGLINAWTAENDRWPI 216
OY 1211 KVDLAPMIRHIGIKTGAROKFSLYISOFIYSLDGKKMQTYRGNSTGLVFEFGND 1270
DB 217 QIMLQRMRYTGVIITGAKRIGSPETIKSYKTIAYSDGKTWAMYKKGTNEEMVRGND 276
OY 1271 SSGIKHNIFNPPIIARYIRLPHYSIRSTLRMLMGCDLNSCMLPMEKSKAIDAOIT 1330
DB 277 NNPVANSFPPRIKQYVRLPQICRRHCTLRRELGLCGLSCGSEPLGMKSGHIDYIT 336
OY 1331 ASSYF----TNMFAIYSPSKARLHLOGSNAMRPOYNNKEKMLQVDFOKTMKVYVTTQ 1386
DB 337 ASSVFTLNNDMF-TWEPKARLDKQGVNMTSGHDSQWLOYDLVPTVGTGIIITQ 395
OY 1387 VSLTSMYVKEFLISSODGHQWTLF--FONGKVVYFQNDSEFTPVVNSLDPPLR 1444
DB 396 AKDFGHVQFVGSFKLYISNDGEHMYHODEKQKDKVFOGDFNDTHRKNVLDPPIYARF 455
OY 1445 LRTHPOSWVHQIALRMEVLGCEAOD 1469
DB 456 IRLPMSWYGRITLRSBLGCAEE 480
```

Search completed: November 17, 2001, 13:25:24  
Job time: 188 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2001, 13:22:16 ; Search time 16.43 seconds  
(without alignments)  
3066.940 Million cell updates/sec

Title: US-09-689-430-2  
Perfect score: 7669  
Sequence: 1 MQIELSTCFELCLIRFCFSA.....WVHQLRMVLEGEADLY 1471

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues  
Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7412	94.2	2351	1	FA8_HUMAN
2	6347	80.7	2133	1	FA8_PIG
3	6301	-80.1	2319	1	FA8_MOUSE
4	2405.5	30.6	2224	1	FA5_HUMAN
5	2379	30.2	2211	1	FA5_BOVIN
6	1725	21.9	1065	1	CERU_HUMAN
7	1678	21.3	1059	1	CERU_RAT
8	1639	20.8	1062	1	CERU_MOUSE
9	663	8.4	427	1	MFGM_RAT
10	657	8.3	463	1	MFGM_MOUSE
11	650	8.3	409	1	MFGM_PIG
12	635	8.1	427	1	MFGM_BOVIN
13	588	7.5	387	1	MFGM_HUMAN
14	469.5	6.0	931	1	MRP2_HUMAN
15	464.5	5.9	925	1	MRP2_RAT
16	462.5	5.9	931	1	MRP2_MOUSE
17	458.5	5.8	914	1	MRP1_CHICK
18	451.5	5.7	922	1	MRP1_RAT
19	446.5	5.7	923	1	MRP1_MOUSE
20	443	5.6	928	1	MRP1_XENLA
21	429.5	5.5	923	1	MRP1_HUMAN
22	306.5	3.3	3133	1	HMCT_BOVMO
23	260.5	3.3	280	1	XR1L_FUGRU
24	247	3.1	224	1	XR1L_MOUSE
25	243	3.1	224	1	XR1L_HUMAN
26	198.5	2.5	854	1	DDR2_MOUSE
27	192	2.4	913	1	DDR1_HUMAN
28	185.5	2.4	855	1	DDR2_HUMAN
29	185.5	2.4	910	1	DDR1_RAT
30	149.5	1.9	911	1	DDR1_MOUSE
31	142	1.8	578	1	ASO_TOBAC
32	135.5	1.7	622	1	YAK8_SCHPO
33	135.5	1.7	1196	1	BXCN_CLOBO

34	131.5	1.7	994	1	DP02_KLULA
35	131.5	1.7	4466	1	DYHC_ANTCR
36	130.5	1.7	2216	1	YCF2_EPRIV
37	129	1.6	818	1	DAP2_YEAST
38	128	1.6	1333	1	SOS1_HUMAN
39	128	1.6	1709	1	CHD1_HUMAN
40	127	1.6	2368	1	ESR1_YEAST
41	126.5	1.6	1648	1	YJ9H_YEAST
42	125.5	1.6	1711	1	CHD1_MOUSE
43	124.5	1.6	2280	1	YCF2_TOBAC
44	124	1.6	965	1	AMPN_MOUSE
45	124	1.6	4196	1	DYHC_SCHPO

## ALIGNMENTS

RESULT	ID	FA8_HUMAN	STANDARD:	PRT:	2351 AA.
AC	P00451				
DT	21-JUL-1986	(Rel. 01, Created)			
DT	21-JUL-1986	(Rel. 01, Last sequence update)			
DT	01-OCT-2000	(Rel. 40, Last annotation update)			
DE	COAGULATION FACTOR VIII PRECURSOR (PROCOAGULANT COMPONENT)				
DE	(ANTHEMOPHILIC FACTOR) (AHF).				
GN	F8 OR F8C.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=86081164; PubMed=3935400;				
RA	Truett M.A., Blacher R., Burke R.L., Caput D., Chu C., Dina D.,				
RA	Harlog K., Kuo C.H., Maslarz F.R., Merryweather J.P., Najjarin A.,				
RA	Pachl C., Potter S.J., Puma J., Quiroga M., Rall L.B., Randolph A.,				
RA	Urdea M.S., Valenzuela P., Dahl H.-H.M., Favalaro J., Hansen J.,				
RA	Norfang O., Ezban M.,				
RT	"Characterization of the polypeptide composition of human factor				
RT	VIII:C and the nucleotide sequence and expression of the human kidney				
RT	cDNA.";				
RL	DNA 4:333-349(1985).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=85061548; PubMed=6438526;				
RA	Wood W.I., Capon D.J., Simonsen C.C., Eaton D.L., Gitschler J.,				
RA	Key B., Seeburg P.H., Smith D.H., Hollingshead P., Wilson K.L.,				
RA	Delwart E., Tuddenham E.G.D., Vehar G.A., Lawn R.M.;				
RT	"Expression of active human factor VIII from recombinant DNA clones.";				
RL	Nature 312:330-337(1984).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=85061550; PubMed=6438528;				
RA	Toole J.J., Knopf J.L., Wozney J.M., Sultzman L.A., Buckner J.L.,				
RA	Pitman D.D., Kaufman R.J., Brown E., Shoemaker C., Orr E.C.,				
RA	Amphlett G.W., Foster W.B., Coe M.L., Knutson G.J., Pass D.N.,				
RA	Hewick R.M.;				
RT	"Molecular cloning of a cDNA encoding human antihemophilic factor.";				
RL	Nature 312:342-347(1984).				
RN	[4]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=93265012; PubMed=1303178;				
RA	Gitschler J., Wood W.I.;				
RT	"Sequence of the exon-containing regions of the human factor VIII				
RL	gene.";				
RL	Hum. Mol. Genet. 1:199-200(1992).				
RN	[5]				
RP	SEQUENCE OF 2064-2070 FROM N.A.				
RA	de Water N.S., Williams R., Browett P.J.;				
RL	Submitted (JUN-1997) to the EMBL/Genbank/DBJ databases.				
RN	[6]				
RP	SULFATATION OF 1699.				



RX MEDLINE-91093266; PubMed-1898735;  
 RA Leyte A., van Schijndel H.B., Niehrs C., Huttner W.B., Verbeet M.P.,  
 RA Mertens K., van Mourik J.A.;  
 RT "Sulfation of Tyr1680 of human blood coagulation factor VIII is  
 RT essential for the interaction of factor VIII with von Willebrand  
 RT factor.";   
 RL J. Biol. Chem. 266:740-746(1991).  
 RN [17]  
 RP SULFATATION.  
 RX MEDLINE-92207952; PubMed-1554716;  
 RA Pittman D.D., Wang J.H., Kaufman R.J.;  
 RT "Identification and functional importance of tyrosine sulfate  
 RT residues within recombinant factor VIII.";   
 RL Biochemistry 31:3315-3325(1992).  
 RN [18]  
 RP STRUCTURE BY NMR OF 2322-2343.  
 RX MEDLINE-95200924; PubMed-7893714;  
 RA Gilbert G.E., Baleja J.D.;  
 RT "Membrane-binding peptide from the C2 domain of factor VIII forms an  
 RT amphipathic structure as determined by NMR spectroscopy.";   
 RL Biochemistry 34:3022-3031(1995).  
 RN [19]  
 RP REVIEW ON MOLECULAR BASIS OF HEMOPHILIA A.  
 RX MEDLINE-91221499; PubMed-1902642;  
 RA Gitschler J.;  
 RT "The molecular basis of hemophilia A.";   
 RL Ann. N.Y. Acad. Sci. 614:89-96(1991).  
 RN [10]  
 RP REVIEW ON MOLECULAR BASIS OF HEMOPHILIA A.  
 RX MEDLINE-89088506; PubMed-2491949;  
 RA White G.C. II, Shoemaker C.B.;  
 RT "Factor VIII gene and hemophilia A.";   
 RL Blood 73:1-12(1989).  
 RN [11]  
 RP REVIEW ON MOLECULAR BASIS OF HEMOPHILIA A.  
 RX MEDLINE-95245332; PubMed-7728145;  
 RA Antonarakis S.E., Kazazian H.H., Tuddenham E.G.D.;  
 RT "Molecular etiology of factor VIII deficiency in hemophilia A.";   
 RL Hum. Mutat. 5:1-22(1995).  
 RN [12]  
 RP VARIANT GLN-2326.  
 RX MEDLINE-86235434; PubMed-3012775;  
 RA Gitschler J., Wood W.I., Shuman M.A., Lawn R.M.;  
 RT "Identification of a missense mutation in the factor VIII gene of a  
 RT mild hemophilic.";   
 RL Science 232:1415-1416(1986).  
 RN [13]  
 RP VARIANT PRO-2135.  
 RX MEDLINE-88096539; PubMed-3122181;  
 RA Levinson B., Janco R.L., Phillips J.A. III, Gitschler J.;  
 RT "A novel missense mutation in the factor VIII gene identified by  
 RT analysis of amplified hemophilia DNA sequences.";   
 RL Nucleic Acids Res. 15:9797-9805(1987).  
 RN [14]  
 RP VARIANT GLN-2228.  
 RX MEDLINE-88191889; PubMed-2833855;  
 RA Youssoufian H., Antonarakis S.E., Bell W., Griffin A.M.,  
 RT "Nonsense and missense mutations in hemophilia A: estimate of the  
 RT relative mutation rate at CG dinucleotides.";   
 RL Am. J. Hum. Genet. 42:718-725(1988).  
 RN [15]  
 RP VARIANT GLY-291.  
 RX MEDLINE-88220354; PubMed-2835904;  
 RA Youssoufian H., Wong C., Aronis S., Platokoukis H., Kazazian H.H. Jr.,  
 RA Antonarakis S.E.;  
 RT "Moderately severe hemophilia A resulting from Glu->Gly substitution  
 RT in exon 7 of the factor VIII gene.";   
 RL Am. J. Hum. Genet. 42:867-871(1988).  
 RN [16]  
 RP VARIANT CYS-1708.  
 RX MEDLINE-89274393; PubMed-2499363;  
 RA O'Brien D.P., Tuddenham E.G.;

RT "Purification and characterization of factor VIII 1,689-Cys: a  
 RT nonfunctional cofactor occurring in a patient with severe hemophilia  
 RT A.";   
 RL Blood 73:2117-2122(1989).  
 RN [17]  
 RP VARIANT CYS-391.  
 RX MEDLINE-90001543; PubMed-2506948;  
 RA Shima M., Ware J., Yoshioka A., Fukui H., Fulcher C.A.;  
 RT "An arginine to cysteine amino acid substitution at a critical  
 RT thrombin cleavage site in a dysfunctional factor VIII molecule.";   
 RL Blood 74:1612-1617(1989).  
 RN [18]  
 RP VARIANT LEU-189.  
 RX MEDLINE-90057680; PubMed-2510835;  
 RA Chan V., Chan T.K., Tong T.M., Todd D.;  
 RT "A novel missense mutation in exon 4 of the factor VIII:C gene  
 RT resulting in moderately severe hemophilia A.";   
 RL Blood 74:2688-2691(1989).  
 RN [19]  
 RP VARIANT LEU-2326.  
 RX MEDLINE-89197216; PubMed-2495245;  
 RA Inaba H., Fujimaki M., Kazazian H.H. Jr., Antonarakis S.E.;  
 RT "Mild hemophilia A resulting from Arg-to-Leu substitution in exon 26  
 RT of the factor VIII gene.";   
 RL Hum. Genet. 81:335-338(1989).  
 RN [20]  
 RP VARIANT HIS-391.  
 RX MEDLINE-89264602; PubMed-2498882;  
 RA Arai M., Inaba H., Higuchi M., Antonarakis S.E., Kazazian H.H. Jr.,  
 RA Fujimaki M., Hoyer L.W.;  
 RT "Direct characterization of factor VIII in plasma: detection of a  
 RT mutation altering a thrombin cleavage site  
 RT (arginine-372->histidine).";   
 RL Proc. Natl. Acad. Sci. U.S.A. 86:4277-4281(1989).  
 RN [21]  
 RP VARIANT CYS-1708.  
 RX MEDLINE-90105723; PubMed-2104766;  
 RA Arai M., Higuchi M., Antonarakis S.E., Kazazian H.H. Jr.,  
 RA Phillips J.A. III, Janco R.L., Hoyer L.W.;  
 RT "Characterization of a thrombin cleavage site mutation (Arg 1689 to  
 RT Cys) in the factor VIII gene of two unrelated patients with  
 RT cross-reacting material-positive hemophilia A.";   
 RL Blood 75:384-389(1990).  
 RN [22]  
 RP VARIANTS GLN-2228 AND LEU-2326.  
 RX MEDLINE-90123183; PubMed-2105106;  
 RA Casula L., Murru S., Pecorara M., Ristaldi M.S., Restagno G.,  
 RA Mancuso G., Morfini M., de Biasi R., Baudo F., Carbonara A.;  
 RT "Recurrent mutations and three novel rearrangements in the factor  
 RT VIII gene of hemophilia A patients of Italian descent.";   
 RL Blood 75:662-670(1990).  
 RN [23]  
 RP VARIANT CYS-391.  
 RX MEDLINE-90329422; PubMed-1973901;  
 RA Patinson J.K., McVey J.H., Boon M., Ajani A., Tuddenham E.G.;  
 RT "CRM+ haemophilia A due to a missense mutation (372->Cys) at the  
 RT internal heavy chain thrombin cleavage site.";   
 RL Br. J. Haematol. 75:73-77(1990).  
 RN [24]  
 RP VARIANTS PHE-1699 AND CYS-1708.  
 RX MEDLINE-90152691; PubMed-2105906;  
 RA Higuchi M., Wong C., Kochan L., Olek K., Aronis S., Kasper C.K.,  
 RA Kazazian H.H., Antonarakis S.E.;  
 RT "Characterization of mutations in the factor VIII gene by direct  
 RT sequencing of amplified genomic DNA.";   
 RL Genomics 6:65-71(1990).  
 RN [25]  
 RP VARIANTS CYS-1728 AND ASP-1941.  
 RX MEDLINE-90169988; PubMed-2106480;  
 RA Traystman M.D., Higuchi M., Kasper C.K., Antonarakis S.E.,  
 RA Kazazian H.H.;  
 RT "Use of denaturing gradient gel electrophoresis to detect point  
 RT mutations in the factor VIII gene.";















OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6 X CBA; TISSUE=Liver;  
 RX MEDLINE=93300511; PubMed=8314577;  
 RA Elder B., Lachich D., Gletscher J.;  
 RT "Sequence of the murine factor VIII cDNA";  
 RL Genomics 16:374-379(1993).  
 CC -1- FUNCTION: FACTOR VIII, ALONG WITH CALCIUM AND PHOSPHOLIPID, ACTS  
 CC AS A COFACTOR FOR FACTOR IXA WHEN IT CONVERTS FACTOR X TO THE  
 CC ACTIVATED FORM, FACTOR XA.  
 CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.  
 CC -1- TISSUE SPECIFICITY: FOUND IN MOST TISSUES.  
 CC -1- SIMILARITY: CONTAINS 3 F5/8 TYPE A DOMAINS; EACH IS COMPOSED OF  
 CC 2 PLASTOCYANIN-LIKE REPEATS.  
 CC -1- SIMILARITY: STRONG, TO COAGULATION FACTOR V.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL: L05573; AAA37385.1; -  
 DR PIR: A47004; A47004.  
 DR HSSP: P00451; ICFG.  
 DR MGD: MGI:88383; F8.  
 DR InterPro: IPR000421; -  
 DR InterPro: IPR001117; -  
 DR Pfam: PF00394; Cu-oxidase; 3.  
 DR Pfam: PF00754; F5\_F8\_type\_C; 2.  
 DR PROSITE: PS00079; MULTICOPPER\_OXIDASE1; 3.  
 DR PROSITE: PS01285; FA58C\_1; 2.  
 DR PROSITE: PS01286; FA58C\_2; 2.  
 KW Blood coagulation; Repeat; Plasma; Acute phase; Calcium;  
 KW Signal: Glycoprotein; Sulfatation.  
 FT SIGNAL 1 19  
 FT CHAIN 20 2319  
 FT DOMAIN 20 349  
 FT DOMAIN 20 199  
 FT DOMAIN 207 349  
 FT DOMAIN 399 730  
 FT DOMAIN 399 573  
 FT DOMAIN 583 730  
 FT DOMAIN 760 1640  
 FT DOMAIN 1683 2008  
 FT DOMAIN 1683 1845  
 FT DOMAIN 1855 2008  
 FT DOMAIN 2008 2156  
 FT DOMAIN 2161 2313  
 FT SITE 391 392  
 FT SITE 391 392  
 FT SITE 759 760  
 FT SITE 1678 1679  
 FT SITE 1324 1325  
 FT SITE 1640 1641  
 FT MOD\_RES 367 367  
 FT MOD\_RES 737 737  
 FT MOD\_RES 738 738  
 FT MOD\_RES 742 742  
 FT MOD\_RES 1669 1669  
 FT MOD\_RES 1687 1687  
 FT MOD\_RES 173 199  
 FT DISULFID 547 573  
 FT DISULFID 1819 1845  
 FT DISULFID 2008 2156  
 FT DISULFID 2161 2313

FT CARBOHYD 61 61 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 233 233 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 259 259 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 423 423 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 601 601 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 880 880 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 958 958 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 1015 1015 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 1022 1022 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 1026 1026 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 1044 1044 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 1076 1076 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 1087 1087 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 1136 1136 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 1161 1161 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 1192 1192 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 1255 1255 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 1268 1268 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 1273 1273 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 1274 1274 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 1302 1302 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 1316 1316 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 1340 1340 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 1378 1378 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 1797 1797 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 2105 2105 N-LINKED (GLCNAC...) (POTENTIAL).  
 SQ SEQUENCE 2319 AA; 266148 MW; FD054DE051DB2A01 CRC64;

Query Match 80.1%; Score 6301; DB 1; Length 2319;  
 Best Local Similarity 53.8%; Pred. No. 0;  
 Matches 1251; Conservative 102; Mismatches 111; Indels 862; Gaps 10;

OY 1 MDELSTCEFLCLRCFCFSATRRYYLGAVELSMYDQSD-LGELPVDAEPFRVKSPPF 59  
 DB 1 MOTALFACFFLFLNFCSSAIRYYLGAVELSMNITQSLVLTDSRFLPRMSTSPF 60  
 OY 60 NTSVYKKTLEVEFVHLEINAKRPPMGLGPTIQAEVYDVTYVITLKNASHPVSJHA 119  
 DB 61 NTSIMKKTVEFVKQDLINAKRPPMGLGPTIWEVDVTYVITLKNASHPVSJHA 120  
 OY 120 VGVSYKASGEAYDQGTQREKEDKVPFGSGSHYYWQYLKENGPASDPLCTYSLS 179  
 DB 121 VGVSYKASGEAYDQGTQREKEDKVPFGSGSHYYWQYLKENGPASDPLCTYSLS 180  
 OY 180 HDVLVLDNSGLGALLVCREGSLAKETQTLHKFTLFAVDEGKSWHSEKNSLMODR 239  
 DB 181 HDVLVLDNSGLGALLVCREGSLAKETQTLHKFTLFAVDEGKSWHSEKNSLMODR 240  
 OY 240 DAASARAMPKMTVNGVYVNRSLPGLIGCHRSYVHWVIGMGTTPYVHSTLEGHTFLVRN 299  
 DB 241 DSASARDMPKMTVNGVYVNRSLPGLIGCHRSYVHWVIGMGTTPYVHSTLEGHTFLVRN 300  
 OY 300 HRASLEISPTIFLTQTLMLDGLFCHTSSHOHOMEXVYVVDSCPREPOLRMK-N 358  
 DB 301 HRASLEISPTIFLTQTLMLDGLFCHTSSHOHOMEXVYVVDSCPREPOLRMK-N 360  
 OY 359 NEAEEDYDDLDESDMVDVRFDDDNPSFIQIRSAKRPKRWVYVYIAAEEDMDYAVL 418  
 DB 361 NEAEEDYDDLDESDMVDVRFDDDNPSFIQIRSAKRPKRWVYVYIAAEEDMDYAVL 418  
 OY 419 LAFDRSYKQYLLNNGPQRIIGRKYKVRPMAYTDEFKTRALIQHESGILGLYGEVD 478  
 DB 419 PTDSDNGSYKQYLLNNGPQRIIGRKYKVRPMAYTDEFKTRALIQHESGILGLYGEVD 478  
 OY 479 TLLIFKNOASRPYNYIPGIDVRLPLSRPLPKYKHLKDEPLPGLGELFKKTYVED 538  
 DB 479 TLLIFKNOASRPYNYIPGIDVRLPLSRPLPKYKHLKDEPLPGLGELFKKTYVED 538  
 OY 539 GPKPSDPLCTRYSSFVNMERDLASGLGPLCYKESVDORGNOIMSDRNVTLFSVF 598  
 DB 539 GPKPSDPLCTRYSSFVNMERDLASGLGPLCYKESVDORGNOIMSDRNVTLFSVF 598



QY	599	DENSMTLEMOIFLPRPAVOULEDEFOASNTMHSINGVPEJSLDLSCLHEVAWYT	658
Db	599	DEMOISMTTEMORFLPAAATQPODGFQASNTMHSINGVPEJSLDLSCLHEVAWMI	658
QY	659	LSIAQDTEFLVFFSGYTFKRMAYEDTLTFEPGSETFVMSMENPGLATILGCHNSDFRN	718
Db	659	LSYGAQDTEFLSIFFGYTFKRMAYEDTLTFEPGSETFVMSMENPGLATILGCHNSDFRK	718
QY	719	RGMTALLKVSSCDKNTGDDYEDYSIAVLLSKNNAIEPRFSQNSRHPSTROKOF--	775
Db	719	RGMTALLKVSSCDKSTSDYEIYEDIPLOLVNENNYIDPRFSQNTNHPNTRKKRFDOS	778
QY	776	-----	775
Db	779	TIPKNDMEKIEPQEEIAEMLKVQSVASVSDMLLGLQSHPTPHGLFLSDQEAIEYAID	838
QY	776	-----	775
Db	839	DHSRNALDSENGSPSKVYQLRPESHSEKLYFTQPGQLBSKNSLETTIEYKKKKLGLQY	898
QY	776	-----	775
Db	899	SSLPNSLMTTLLSDNLKATPEKTDSSGFPDMPYHSSSKLSTTAFGKKAVALGVSHVPLN	958
QY	776	-----	775
Db	959	ASEBNSDNLSDTYLWYSOESLPRNDTLSENDRLREKRRFGIALTLTKDNTLFKDNVSL	1011
QY	776	-----	775
Db	1019	MKTNKTINHSTTNKLEKLTBESPTSIENSTTDLQALIKVNSEIOEYVALIHDGTLGKNST	107
QY	776	-----	775
Db	1079	YLRNLHMLNRTSTKKNKDIFPRKDEDPIDQDEENTIMPFSKMLFLSESSNWFKKTGNNS	113
QY	776	-----	775
Db	1139	LNSBOEHSPKOYLVMERKYYKNOSFLSEKNKYVEBODGFTKNIGLKDMAFPHNMSIFLT	119
QY	776	-----	775
Db	1199	TLSNVHENGHNQKNIQOELEKALLEEKVYLPQVHEATGSKNFLKILLTTRONISL	125
QY	776	-----	778
Db	1259	YEVHVPLYQNTSINNSTNYQIIMEHFPRKDKDETNSGLVKNKTRPMYKKNPQOKNT	131
QY	779	-----	778
Db	1319	TQSRKRALGQFRLSTQWLKTIINCSTOCIIQIDHSKEMKFKITKSLSDSVIKSTQTN	137
QY	779	-----	778
Db	1379	SSDSHVKTSAPFPIIDLKRSBPQONKFSHVQASSYIDFKTKSSRIOESNNFLKTRKINP	143
QY	779	-----	778
Db	1439	SLALLPNMFIQDKFTSPGKSNSTNSYTKKRENIIFLKTLEBESGKIELLPQVSIQOE	149
QY	779	-----	778
Db	1499	EILPTEHSGSPGHLINLMEVFLQIGQPIKWNKAKRHGESIKGTESSKNTRSKLLNH	155
QY	779	-----	778
Db	1559	AMDYHYAAQIPKDMWKSKEKSPETIISKQEDTJLSLRPHGNSHSIGANEKONNPQRETTW	161
QY	779	-----	825
Db	1619	VKGQOTQRTCSQIPRYVKLRHQREL--SAPSOEADYDIALITIE-TIEDPDIYSDIKQ	167
QY	826	SPRSFOKTRHRYTIAVERLMDYGSSSPVULNRNRAOSGSPQFKVYFQEFITDGSFTQF	885

Db	1676	GRPSFOQTRHYFLAAVERLMDYGMST8 - HVLLNRRQSDNVDPQFKVYVDFEPTDSSFSOP	1734
Qy	886	LYRSELNHLGLLPPTYRAVEDNIWMTFENQASRPYSFSSLTISEEDROGCAEPRKRF	945
Db	1735	LYRELNHLGLLPPTYRAVEDNIWMTFENQASRPYSFSSLTISYKEDQR - GEEPRRNF	1799
Qy	946	VKPMETKTYEMKVOHNHMAPTKDEFDCKAMAYFSVDVLEKDVHSLGLPILVCHTNTLPA	1005
Db	1794	VKPMETKTYEMKVOHNHMAPTKDEFDCKAMAYFSVDVLEKDVHSLGLPILVCHTNTLPA	1855
Qy	1006	HGRQVTVQVEFALFTTIDETKSMYTFENMRNRACQNIQMEDPTFKENYRPHALNGITM	1065
Db	1854	HGRQVTVQVEFALFTTIDETKSMYTFENMRNRACQNIQMEDPTFKENYRPHALNGITM	1913
Qy	1066	DTLGLVMAOQRIRWYLLTMSGSNENIHSHFSGHYFYRRKKEEYKMALYNLYPCVFETV	1125
Db	1914	DTLGLVMAOQRIRWYLLTMSGSNENIHSHFSGHYFYRRKKEEYKMALYNLYPCVFETL	1973
Qy	1126	EMLPKSGIWEVECELGIEHLHAGNSTLELVYSKQCTPLGMSAGHTRDFQITASGOYQW	1185
Db	1974	EMLPKSGIWEVECELGIEHLHAGNSTLELVYSKQCTPLGMSAGHTRDFQITASGHYQW	2033
Qy	1186	APKLARLHSGSIAMASTKPEFSMIKVDLAPMIIINGITQCARQKFSLLYSQITMYS	1245
Db	2034	APKLARLHSGSIAMASTKPEFSMIKVDLAPMIIINGITQCARQKFSLLYSQITMYS	2093
Qy	1246	LDGKKQDVRNCSGTLMVFFGVNDSSGICKHNIFNPPIARYRLRHPYHSIRSLRML	1305
Db	2094	LDGKKWLSYQGNSTGTLMTFFGVNDSSGICKHNSFNPIIARYRLRHPYHSIRSLRML	2155
Qy	1306	MGCDLNSCSMPLGNEKSAISDAOITASYSYFTNNPATWSPSKARLHLQGRSNMARPQVNP	1365
Db	2154	MGCDLNSCSIPLEGWESKVISDQITASYSYFTNNPATWSPSKARLHLQGRSNMARPQVNP	2213
Qy	1366	KEMLOVDFQTKMYGTGTTGGVGSLSLTSMYVKEFLISSODGHOVTLFFQNKVAVYFQGN	1425
Db	2214	KQWLOVDFQTKMYGTGTTGGVGSLSLTSMYVKEFLISSODGHNHTQLLYNKGAVYFQGN	2273
Qy	1426	QDSFTPVVNSLDPPLRLFRYLRIHPQSVWVHIOALRMELVISCBAQODY 1471	
Db	2274	QDSFTPVVNSLDPPLRLFRYLRIHPQSVWVHIOALRMELVISCBAQODY 2319	

RESULT 4

FA5\_HUMAN STANDARD; PRT; 2224 AA.

AC P12259; Q14285; 01-OCT-1989 (Rel. 12, Created)

DT 01-FEB-1994 (Rel. 28, Last sequence update)

DT 01-OCT-2000 (Rel. 40, Last annotation update)

DE COAGULATION FACTOR V PRECURSOR (ACTIVATED PROTEIN C COFACTOR).

GN Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606; [1]

RP SEQUENCE FROM N.A.

RA MEDLINE=92232668; PubMed=1567832; Cripe L.D., Moore K.D., Kane W.H.; "Structure of the gene for human coagulation factor V."; Biochemistry 31:3777-3785(1992). [2]

RP SEQUENCE FROM N.A.

RA MEDLINE=87260886; PubMed=3110773; Jenny R.J., Pittman D.D., Toole J.J., Kriz R.W., Aldape R.A., Hewick R.M., Kaufman R.J., Mann K.G.; "Complete cDNA and derived amino acid sequence of human factor V."; Proc. Natl. Acad. Sci. U.S.A. 84:4846-4850(1987). [3]

RP SEQUENCE OF 1-1600 FROM N.A.

RA MEDLINE=88107560; PubMed=2827731;



RA Kane W.H., Ichinose A., Hagen F.S., Davie E.W.;  
 RT "Cloning of cDNAs coding for the heavy chain region and connecting  
 RT region of human factor V, a blood coagulation factor with four types  
 RT of internal repeats.";  
 RL Biochemistry 26:6508-6514(1987).  
 RN [4]  
 RP SEQUENCE OF 1188-1215 AND 1315-2224 FROM N.A.  
 RX MEDLINE=86313665; PubMed=30922220;  
 RA Kane W.H., Davie E.W.;  
 RT "Cloning of a cDNA coding for human factor V, a blood coagulation  
 RT factor homologous to factor VIII and ceruloplasmin.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 83:6800-6804(1986).  
 RN [5]  
 RP PARTIAL SEQUENCE FROM N.A.  
 RC TISSUE=Fibrinolytic;  
 RX MEDLINE=93203619; PubMed=8454869;  
 RA Shen N.L.L., Fan S.-T., Pyatt J., Graff R., Lapolla R.J.,  
 RA Edgington T.S.;  
 RT "The serine protease cofactor factor V is synthesized by  
 RT lymphocytes.";  
 RL J. Immunol. 150:2992-3001(1993).  
 RN [6]  
 RP VARIANT APCR GLN-534.  
 RX MEDLINE=94217810; PubMed=8164741;  
 RA Bertina R.M., Koelman B.P.C., Koster T., Rosendaal F.R.,  
 RA Driven R.J., de Ronde H., van der Velde P.A., Reitsma P.H.;  
 RT "Mutation in blood coagulation factor V associated with resistance to  
 RT activated protein C.";  
 RL Nature 369:64-67(1994).  
 CC -1- FUNCTION: COAGULATION FACTOR V IS A COFACTOR THAT PARTICIPATES  
 CC WITH FACTOR XA TO ACTIVATE PROTHROMBIN TO THROMBIN.  
 CC -1- SUBUNIT: FACTOR VA IS COMPOSED OF AN HEAVY CHAIN AND OF A LIGHT  
 CC CHAIN NONCOVALENTLY BOUND. THE INTERACTION BETWEEN THE TWO CHAINS  
 CC IS CALCIUM-DEPENDENT.  
 CC -1- DOMAIN: DOMAIN B CONTAINS 35 X 9 AA TANDEM REPEATS, AND 2 X 17 AA  
 CC REPEATS.  
 CC -1- PRIM: THROMBIN ACTIVATES FACTOR V PROTEOLYTICALLY TO THE ACTIVE  
 CC COFACTOR, FACTOR V(A) (FORMATION OF A HEAVY CHAIN AT THE N-  
 CC TERMINUS AND A LIGHT CHAIN AT THE C-TERMINUS).  
 CC -1- DISEASE: OMEN PARAHEMOPHILIA, AN HEMORRHAGIC DIATHESIS, IS DUE  
 CC TO A DEFICIENCY OF FACTOR V. OTHER DEFECTS IN F5 RESULTS IN A  
 CC FORM OF THROMBOPHILIA KNOWN AS APC RESISTANCE (APCR). THE APCR  
 CC MUTATION IS FOUND IN ABOUT 5% OF THE POPULATION WHICH SUGGESTS THAT  
 CC A SLIGHT THROMBOTIC TENDENCY MAY CONFER SOME ADVANTAGE IN FETAL  
 CC IMPLANTATION.  
 CC -1- SIMILARITY: CONTAINS 3 F5/8 TYPE A DOMAINS; EACH IS COMPOSED OF  
 CC 2 PLASTOCYANIN-LIKE REPEATS.  
 CC -1- SIMILARITY: STRONG, TO COAGULATION FACTOR VIII.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sb.ch/announce/>  
 CC or send an email to [license@isb-sb.ch](mailto:license@isb-sb.ch)).  
 CC -----  
 DR EMBL: L32779; AAB59401.1; JOINED.  
 DR EMBL: L32755; AAB59401.1; JOINED.  
 DR EMBL: L32756; AAB59401.1; JOINED.  
 DR EMBL: L32757; AAB59401.1; JOINED.  
 DR EMBL: L32758; AAB59401.1; JOINED.  
 DR EMBL: L32759; AAB59401.1; JOINED.  
 DR EMBL: L32760; AAB59401.1; JOINED.  
 DR EMBL: L32761; AAB59401.1; JOINED.  
 DR EMBL: L32762; AAB59401.1; JOINED.  
 DR EMBL: L32763; AAB59401.1; JOINED.  
 DR EMBL: L32764; AAB59401.1; JOINED.  
 DR EMBL: L32765; AAB59401.1; JOINED.  
 DR EMBL: L32766; AAB59401.1; JOINED.  
 DR EMBL: L32767; AAB59401.1; JOINED.  
 DR EMBL: L32768; AAB59401.1; JOINED.

DR EMBL: L32769; AAB59401.1; JOINED.  
 DR EMBL: L32770; AAB59401.1; JOINED.  
 DR EMBL: L32771; AAB59401.1; JOINED.  
 DR EMBL: L32772; AAB59401.1; JOINED.  
 DR EMBL: L32773; AAB59401.1; JOINED.  
 DR EMBL: L32774; AAB59401.1; JOINED.  
 DR EMBL: L32775; AAB59401.1; JOINED.  
 DR EMBL: L32776; AAB59401.1; JOINED.  
 DR EMBL: L32777; AAB59401.1; JOINED.  
 DR EMBL: L32778; AAB59401.1; JOINED.  
 DR EMBL: L32779; AAB59401.1; JOINED.  
 DR EMBL: L32780; AAB59401.1; JOINED.  
 DR EMBL: L32781; AAB59401.1; JOINED.  
 DR EMBL: L32782; AAB59401.1; JOINED.  
 DR EMBL: L32783; AAB59401.1; JOINED.  
 DR EMBL: L32784; AAB59401.1; JOINED.  
 DR EMBL: L32785; AAB59401.1; JOINED.  
 DR EMBL: L32786; AAB59401.1; JOINED.  
 DR EMBL: L32787; AAB59401.1; JOINED.  
 DR EMBL: L32788; AAB59401.1; JOINED.  
 DR EMBL: L32789; AAB59401.1; JOINED.  
 DR EMBL: L32790; AAB59401.1; JOINED.  
 DR EMBL: L32791; AAB59401.1; JOINED.  
 DR EMBL: L32792; AAB59401.1; JOINED.  
 DR EMBL: L32793; AAB59401.1; JOINED.  
 DR EMBL: L32794; AAB59401.1; JOINED.  
 DR EMBL: L32795; AAB59401.1; JOINED.  
 DR EMBL: L32796; AAB59401.1; JOINED.  
 DR EMBL: L32797; AAB59401.1; JOINED.  
 DR EMBL: L32798; AAB59401.1; JOINED.  
 DR EMBL: L32799; AAB59401.1; JOINED.  
 DR EMBL: L32800; AAB59401.1; JOINED.  
 DR EMBL: L32801; AAB59401.1; JOINED.  
 DR EMBL: L32802; AAB59401.1; JOINED.  
 DR EMBL: L32803; AAB59401.1; JOINED.  
 DR EMBL: L32804; AAB59401.1; JOINED.  
 DR EMBL: L32805; AAB59401.1; JOINED.  
 DR EMBL: L32806; AAB59401.1; JOINED.  
 DR EMBL: L32807; AAB59401.1; JOINED.  
 DR EMBL: L32808; AAB59401.1; JOINED.  
 DR EMBL: L32809; AAB59401.1; JOINED.  
 DR EMBL: L32810; AAB59401.1; JOINED.  
 DR EMBL: L32811; AAB59401.1; JOINED.  
 DR EMBL: L32812; AAB59401.1; JOINED.  
 DR EMBL: L32813; AAB59401.1; JOINED.  
 DR EMBL: L32814; AAB59401.1; JOINED.  
 DR EMBL: L32815; AAB59401.1; JOINED.  
 DR EMBL: L32816; AAB59401.1; JOINED.  
 DR EMBL: L32817; AAB59401.1; JOINED.  
 DR EMBL: L32818; AAB59401.1; JOINED.  
 DR EMBL: L32819; AAB59401.1; JOINED.  
 DR EMBL: L32820; AAB59401.1; JOINED.  
 DR EMBL: L32821; AAB59401.1; JOINED.  
 DR EMBL: L32822; AAB59401.1; JOINED.  
 DR EMBL: L32823; AAB59401.1; JOINED.  
 DR EMBL: L32824; AAB59401.1; JOINED.  
 DR EMBL: L32825; AAB59401.1; JOINED.  
 DR EMBL: L32826; AAB59401.1; JOINED.  
 DR EMBL: L32827; AAB59401.1; JOINED.  
 DR EMBL: L32828; AAB59401.1; JOINED.  
 DR EMBL: L32829; AAB59401.1; JOINED.  
 DR EMBL: L32830; AAB59401.1; JOINED.  
 DR EMBL: L32831; AAB59401.1; JOINED.  
 DR EMBL: L32832; AAB59401.1; JOINED.  
 DR EMBL: L32833; AAB59401.1; JOINED.  
 DR EMBL: L32834; AAB59401.1; JOINED.  
 DR EMBL: L32835; AAB59401.1; JOINED.  
 DR EMBL: L32836; AAB59401.1; JOINED.  
 DR EMBL: L32837; AAB59401.1; JOINED.  
 DR EMBL: L32838; AAB59401.1; JOINED.  
 DR EMBL: L32839; AAB59401.1; JOINED.  
 DR EMBL: L32840; AAB59401.1; JOINED.  
 DR EMBL: L32841; AAB59401.1; JOINED.  
 DR EMBL: L32842; AAB59401.1; JOINED.  
 DR EMBL: L32843; AAB59401.1; JOINED.  
 DR EMBL: L32844; AAB59401.1; JOINED.  
 DR EMBL: L32845; AAB59401.1; JOINED.  
 DR EMBL: L32846; AAB59401.1; JOINED.  
 DR EMBL: L32847; AAB59401.1; JOINED.  
 DR EMBL: L32848; AAB59401.1; JOINED.  
 DR EMBL: L32849; AAB59401.1; JOINED.  
 DR EMBL: L32850; AAB59401.1; JOINED.  
 DR EMBL: L32851; AAB59401.1; JOINED.  
 DR EMBL: L32852; AAB59401.1; JOINED.  
 DR EMBL: L32853; AAB59401.1; JOINED.  
 DR EMBL: L32854; AAB59401.1; JOINED.  
 DR EMBL: L32855; AAB59401.1; JOINED.  
 DR EMBL: L32856; AAB59401.1; JOINED.  
 DR EMBL: L32857; AAB59401.1; JOINED.  
 DR EMBL: L32858; AAB59401.1; JOINED.  
 DR EMBL: L32859; AAB59401.1; JOINED.  
 DR EMBL: L32860; AAB59401.1; JOINED.  
 DR EMBL: L32861; AAB59401.1; JOINED.  
 DR EMBL: L32862; AAB59401.1; JOINED.  
 DR EMBL: L32863; AAB59401.1; JOINED.  
 DR EMBL: L32864; AAB59401.1; JOINED.  
 DR EMBL: L32865; AAB59401.1; JOINED.  
 DR EMBL: L32866; AAB59401.1; JOINED.  
 DR EMBL: L32867; AAB59401.1; JOINED.  
 DR EMBL: L32868; AAB59401.1; JOINED.  
 DR EMBL: L32869; AAB59401.1; JOINED.  
 DR EMBL: L32870; AAB59401.1; JOINED.  
 DR EMBL: L32871; AAB59401.1; JOINED.  
 DR EMBL: L32872; AAB59401.1; JOINED.  
 DR EMBL: L32873; AAB59401.1; JOINED.  
 DR EMBL: L32874; AAB59401.1; JOINED.  
 DR EMBL: L32875; AAB59401.1; JOINED.  
 DR EMBL: L32876; AAB59401.1; JOINED.  
 DR EMBL: L32877; AAB59401.1; JOINED.  
 DR EMBL: L32878; AAB59401.1; JOINED.  
 DR EMBL: L32879; AAB59401.1; JOINED.  
 DR EMBL: L32880; AAB59401.1; JOINED.  
 DR EMBL: L32881; AAB59401.1; JOINED.  
 DR EMBL: L32882; AAB59401.1; JOINED.  
 DR EMBL: L32883; AAB59401.1; JOINED.  
 DR EMBL: L32884; AAB59401.1; JOINED.  
 DR EMBL: L32885; AAB59401.1; JOINED.  
 DR EMBL: L32886; AAB59401.1; JOINED.  
 DR EMBL: L32887; AAB59401.1; JOINED.  
 DR EMBL: L32888; AAB59401.1; JOINED.  
 DR EMBL: L32889; AAB59401.1; JOINED.  
 DR EMBL: L32890; AAB59401.1; JOINED.  
 DR EMBL: L32891; AAB59401.1; JOINED.  
 DR EMBL: L32892; AAB59401.1; JOINED.  
 DR EMBL: L32893; AAB59401.1; JOINED.  
 DR EMBL: L32894; AAB59401.1; JOINED.  
 DR EMBL: L32895; AAB59401.1; JOINED.  
 DR EMBL: L32896; AAB59401.1; JOINED.  
 DR EMBL: L32897; AAB59401.1; JOINED.  
 DR EMBL: L32898; AAB59401.1; JOINED.  
 DR EMBL: L32899; AAB59401.1; JOINED.  
 DR EMBL: L32900; AAB59401.1; JOINED.  
 DR EMBL: L32901; AAB59401.1; JOINED.  
 DR EMBL: L32902; AAB59401.1; JOINED.  
 DR EMBL: L32903; AAB59401.1; JOINED.  
 DR EMBL: L32904; AAB59401.1; JOINED.  
 DR EMBL: L32905; AAB59401.1; JOINED.  
 DR EMBL: L32906; AAB59401.1; JOINED.  
 DR EMBL: L32907; AAB59401.1; JOINED.  
 DR EMBL: L32908; AAB59401.1; JOINED.  
 DR EMBL: L32909; AAB59401.1; JOINED.  
 DR EMBL: L32910; AAB59401.1; JOINED.  
 DR EMBL: L32911; AAB59401.1; JOINED.  
 DR EMBL: L32912; AAB59401.1; JOINED.  
 DR EMBL: L32913; AAB59401.1; JOINED.  
 DR EMBL: L32914; AAB59401.1; JOINED.  
 DR EMBL: L32915; AAB59401.1; JOINED.  
 DR EMBL: L32916; AAB59401.1; JOINED.  
 DR EMBL: L32917; AAB59401.1; JOINED.  
 DR EMBL: L32918; AAB59401.1; JOINED.  
 DR EMBL: L32919; AAB59401.1; JOINED.  
 DR EMBL: L32920; AAB59401.1; JOINED.  
 DR EMBL: L32921; AAB59401.1; JOINED.  
 DR EMBL: L32922; AAB59401.1; JOINED.  
 DR EMBL: L32923; AAB59401.1; JOINED.  
 DR EMBL: L32924; AAB59401.1; JOINED.  
 DR EMBL: L32925; AAB59401.1; JOINED.  
 DR EMBL: L32926; AAB59401.1; JOINED.  
 DR EMBL: L32927; AAB59401.1; JOINED.  
 DR EMBL: L32928; AAB59401.1; JOINED.  
 DR EMBL: L32929; AAB59401.1; JOINED.  
 DR EMBL: L32930; AAB59401.1; JOINED.  
 DR EMBL: L32931; AAB59401.1; JOINED.  
 DR EMBL: L32932; AAB59401.1; JOINED.  
 DR EMBL: L32933; AAB59401.1; JOINED.  
 DR EMBL: L32934; AAB59401.1; JOINED.  
 DR EMBL: L32935; AAB59401.1; JOINED.  
 DR EMBL: L32936; AAB59401.1; JOINED.  
 DR EMBL: L32937; AAB59401.1; JOINED.  
 DR EMBL: L32938; AAB59401.1; JOINED.  
 DR EMBL: L32939; AAB59401.1; JOINED.  
 DR EMBL: L32940; AAB59401.1; JOINED.  
 DR EMBL: L32941; AAB59401.1; JOINED.  
 DR EMBL: L32942; AAB59401.1; JOINED.  
 DR EMBL: L32943; AAB59401.1; JOINED.  
 DR EMBL: L32944; AAB59401.1; JOINED.  
 DR EMBL: L32945; AAB59401.1; JOINED.  
 DR EMBL: L32946; AAB59401.1; JOINED.  
 DR EMBL: L32947; AAB59401.1; JOINED.  
 DR EMBL: L32948; AAB59401.1; JOINED.  
 DR EMBL: L32949; AAB59401.1; JOINED.  
 DR EMBL: L32950; AAB59401.1; JOINED.  
 DR EMBL: L32951; AAB59401.1; JOINED.  
 DR EMBL: L32952; AAB59401.1; JOINED.  
 DR EMBL: L32953; AAB59401.1; JOINED.  
 DR EMBL: L32954; AAB59401.1; JOINED.  
 DR EMBL: L32955; AAB59401.1; JOINED.  
 DR EMBL: L32956; AAB59401.1; JOINED.  
 DR EMBL: L32957; AAB59401.1; JOINED.  
 DR EMBL: L32958; AAB59401.1; JOINED.  
 DR EMBL: L32959; AAB59401.1; JOINED.  
 DR EMBL: L32960; AAB59401.1; JOINED.  
 DR EMBL: L32961; AAB59401.1; JOINED.  
 DR EMBL: L32962; AAB59401.1; JOINED.  
 DR EMBL: L32963; AAB59401.1; JOINED.  
 DR EMBL: L32964; AAB59401.1; JOINED.  
 DR EMBL: L32965; AAB59401.1; JOINED.  
 DR EMBL: L32966; AAB59401.1; JOINED.  
 DR EMBL: L32967; AAB59401.1; JOINED.  
 DR EMBL: L32968; AAB59401.1; JOINED.  
 DR EMBL: L32969; AAB59401.1; JOINED.  
 DR EMBL: L32970; AAB59401.1; JOINED.  
 DR EMBL: L32971; AAB59401.1; JOINED.  
 DR EMBL: L32972; AAB59401.1; JOINED.  
 DR EMBL: L32973; AAB59401.1; JOINED.  
 DR EMBL: L32974; AAB59401.1; JOINED.  
 DR EMBL: L32975; AAB59401.1; JOINED.  
 DR EMBL: L32976; AAB59401.1; JOINED.  
 DR EMBL: L32977; AAB59401.1; JOINED.  
 DR EMBL: L32978; AAB59401.1; JOINED.  
 DR EMBL: L32979; AAB59401.1; JOINED.  
 DR EMBL: L32980; AAB59401.1; JOINED.  
 DR EMBL: L32981; AAB59401.1; JOINED.  
 DR EMBL: L32982; AAB59401.1; JOINED.  
 DR EMBL: L32983; AAB59401.1; JOINED.  
 DR EMBL: L32984; AAB59401.1; JOINED.  
 DR EMBL: L32985; AAB59401.1; JOINED.  
 DR EMBL: L32986; AAB59401.1; JOINED.  
 DR EMBL: L32987; AAB59401.1; JOINED.  
 DR EMBL: L32988; AAB59401.1; JOINED.  
 DR EMBL: L32989; AAB59401.1; JOINED.  
 DR EMBL: L32990; AAB59401.1; JOINED.  
 DR EMBL: L32991; AAB59401.1; JOINED.  
 DR EMBL: L32992; AAB59401.1; JOINED.  
 DR EMBL: L32993; AAB59401.1; JOINED.  
 DR EMBL: L32994; AAB59401.1; JOINED.  
 DR EMBL: L32995; AAB59401.1; JOINED.  
 DR EMBL: L32996; AAB59401.1; JOINED.  
 DR EMBL: L32997; AAB59401.1; JOINED.  
 DR EMBL: L32998; AAB59401.1; JOINED.  
 DR EMBL: L32999; AAB59401.1; JOINED.  
 DR EMBL: L33000; AAB59401.1; JOINED.  
 DR EMBL: L33001; AAB59401.1; JOINED.  
 DR EMBL: L33002; AAB59401.1; JOINED.  
 DR EMBL: L33003; AAB59401.1; JOINED.  
 DR EMBL: L33004; AAB59401.1; JOINED.  
 DR EMBL: L33005; AAB59401.1; JOINED.  
 DR EMBL: L33006; AAB59401.1; JOINED.  
 DR EMBL: L33007; AAB59401.1; JOINED.  
 DR EMBL: L33008; AAB59401.1; JOINED.  
 DR EMBL: L33009; AAB59401.1; JOINED.  
 DR EMBL: L33010; AAB59401.1; JOINED.  
 DR EMBL: L33011; AAB59401.1; JOINED.  
 DR EMBL: L33012; AAB59401.1; JOINED.  
 DR EMBL: L33013; AAB59401.1; JOINED.  
 DR EMBL: L33014; AAB59401.1; JOINED.  
 DR EMBL: L33015; AAB59401.1; JOINED.  
 DR EMBL: L33016; AAB59401.1; JOINED.  
 DR EMBL: L33017; AAB59401.1; JOINED.  
 DR EMBL: L33018; AAB59401.1; JOINED.  
 DR EMBL: L33019; AAB59401.1; JOINED.  
 DR EMBL: L33020; AAB59401.1; JOINED.  
 DR EMBL: L33021; AAB59401.1; JOINED.  
 DR EMBL: L33022; AAB59401.1; JOINED.  
 DR EMBL: L33023; AAB59401.1; JOINED.  
 DR EMBL: L33024; AAB59401.1; JOINED.  
 DR EMBL: L33025; AAB59401.1; JOINED.  
 DR EMBL: L33026; AAB59401.1; JOINED.  
 DR EMBL: L33027; AAB59401.1; JOINED.  
 DR EMBL: L33028; AAB59401.1; JOINED.  
 DR EMBL: L33029; AAB59401.1; JOINED.  
 DR EMBL: L33030; AAB59401.1; JOINED.  
 DR EMBL: L33031; AAB59401.1; JOINED.  
 DR EMBL: L33032; AAB59401.1; JOINED.  
 DR EMBL: L33033; AAB59401.1; JOINED.  
 DR EMBL: L33034; AAB59401.1; JOINED.  
 DR EMBL: L33035; AAB59401.1; JOINED.  
 DR EMBL: L33036; AAB59401.1; JOINED.  
 DR EMBL: L33037; AAB59401.1; JOINED.  
 DR EMBL: L33038; AAB59401.1; JOINED.  
 DR EMBL: L33039; AAB59401.1; JOINED.  
 DR EMBL: L33040; AAB59401.1; JOINED.  
 DR EMBL: L33041; AAB59401.1; JOINED.  
 DR EMBL: L33042; AAB59401.1; JOINED.  
 DR EMBL: L33043; AAB59401.1; JOINED.  
 DR EMBL: L33044; AAB59401.1; JOINED.  
 DR EMBL: L33045; AAB59401.1; JOINED.  
 DR EMBL: L33046; AAB59401.1; JOINED.  
 DR EMBL: L33047; AAB59401.1; JOINED.  
 DR EMBL: L33048; AAB59401.1; JOINED.  
 DR EMBL: L33049; AAB59401.1; JOINED.  
 DR EMBL: L33050; AAB59401.1; JOINED.  
 DR EMBL: L33051; AAB59401.1; JOINED.  
 DR EMBL: L33052; AAB59401.1; JOINED.  
 DR EMBL: L33053; AAB59401.1; JOINED.  
 DR EMBL: L33054; AAB59401.1; JOINED.  
 DR EMBL: L33055; AAB59401.1; JOINED.  
 DR EMBL: L33056; AAB59401.1; JOINED.  
 DR EMBL: L33057; AAB59401.1; JOINED.  
 DR EMBL: L33058; AAB59401.1; JOINED.  
 DR EMBL: L33059; AAB59401.1; JOINED.  
 DR EMBL: L33060; AAB59401.1; JOINED.  
 DR EMBL: L33061; AAB59401.1; JOINED.  
 DR EMBL: L33062; AAB59401.1; JOINED.  
 DR EMBL: L33063; AAB59401.1; JOINED.  
 DR EMBL: L33064; AAB59401.1; JOINED.  
 DR EMBL: L33065; AAB59401.1; JOINED.  
 DR EMBL: L33066; AAB59401.1; JOINED.  
 DR EMBL: L33067; AAB59401.1; JOINED.  
 DR EMBL: L33068; AAB59401.1; JOINED.  
 DR EMBL: L33069; AAB59401.1; JOINED.  
 DR EMBL: L33070; AAB59401.1; JOINED.  
 DR EMBL: L33071; AAB59401.1; JOINED.  
 DR EMBL: L33072; AAB59401.1; JOINED.  
 DR EMBL: L33073; AAB59401.1; JOINED.  
 DR EMBL: L33074; AAB59401.1; JOINED.  
 DR EMBL: L33075; AAB59401.1; JOINED.  
 DR EMBL: L33076; AAB59401.1; JOINED.  
 DR EMBL: L33077; AAB59401.1; JOINED.  
 DR EMBL: L33078; AAB59401.1; JOINED.  
 DR EMBL: L33079; AAB59401.1; JOINED.  
 DR EMBL: L33080; AAB59401.1; JOINED.  
 DR EMBL: L33081; AAB59401.1; JOINED.  
 DR EMBL: L33082; AAB59401.1; JOINED.  
 DR EMBL: L33083; AAB59401.1; JOINED.  
 DR EMBL: L33084; AAB59401.1; JOINED.  
 DR EMBL: L33085; AAB59401.1; JOINED.  
 DR EMBL: L33086; AAB59401.1; JOINED.  
 DR EMBL: L33087; AAB59401.1; JOINED.  
 DR EMBL: L33088; AAB59401.1; JOINED.  
 DR EMBL: L33089; AAB59401.1; JOINED.  
 DR EMBL: L33090; AAB59401.1; JOINED.  
 DR EMBL: L33091; AAB59401.1; JOINED.  
 DR EMBL: L33092; AAB59401.1; JOINED.  
 DR EMBL: L33093; AAB59401.1; JOINED.  
 DR EMBL: L33094; AAB59401.1; JOINED.  
 DR EMBL: L33095; AAB59401.1; JOINED.  
 DR EMBL: L33096; AAB59401.1; JOINED.  
 DR EMBL: L33097; AAB59401.1; JOINED.  
 DR EMBL: L33098; AAB59401.1; JOINED.  
 DR EMBL: L33099; AAB59401.1; JOINED.  
 DR EMBL: L33100; AAB59401.1; JOINED.  
 DR EMBL: L33101; AAB59401.1; JOINED.  
 DR EMBL: L33102; AAB59401.1; JOINED.  
 DR EMBL: L33103; AAB59401.1; JOINED.  
 DR EMBL: L33104; AAB59401.1; JOINED.  
 DR EMBL: L33105; AAB59401.1; JOINED.  
 DR EMBL: L33106; AAB59401.1; JOINED.  
 DR EMBL: L33107; AAB59401.1; JOINED.  
 DR EMBL: L33108; AAB59401.1; JOINED.  
 DR EMBL: L33109; AAB59401.1; JOINED.  
 DR EMBL: L33110; AAB59401.1; JOINED.  
 DR EMBL: L33111; AAB59401.1; JOINED.  
 DR EMBL: L33112; AAB59401.1; JOINED.  
 DR EMBL: L33113; AAB59401.1; JOINED.  
 DR EMBL: L33114; AAB59401.1; JOINED.  
 DR EMBL: L33115; AAB59401.1; JOINED.  
 DR EMBL: L33116; AAB59401.1; JOINED.  
 DR EMBL: L33117; AAB59401.1; JOINED.  
 DR EMBL: L33118; AAB59401.1; JOINED.  
 DR EMBL: L33119; AAB59401.1; JOINED.  
 DR EMBL: L33120; AAB59401.1; JOINED.  
 DR EMBL: L33121; AAB59401.1; JOINED.  
 DR EMBL: L33122; AAB59401.1; JOINED.  
 DR EMBL: L33123; AAB59401.1; JOINED.  
 DR EMBL: L33124; AAB59401.1; JOINED.  
 DR EMBL: L33125; AAB59401.1; JOINED.  
 DR EMBL: L33126; AAB59401.1; JOINED.  
 DR EMBL: L33127; AAB59401.1; JOINED.  
 DR EMBL: L33128; AAB59401.1; JOINED.  
 DR EMBL: L33129; AAB59401.1; JOINED.  
 DR EMBL: L33130; AAB59401.1; JOINED.  
 DR EMBL: L33131; AAB59401.1; JOINED.  
 DR EMBL:







Db	1639	ENHGLICPLTIRAEVDVYIQVRFKMLASPRYSLHAHGLSTYKSSSEKITYEDDSSPEMFKFDN	1639
Qy	945	EVKPNERTKTYFMKVOQHMAPRTKDEDFDCKAMAYFSDVDLEKDVHSGLLGPLVCHTNTLNP	1004
Db	1699	AVQPNSSYTYVMHATERSGSPSPSACRAMAYSAVNPENKQIHSLGILLOKQGLIHK	1756
Qy	1005	AHGROYTVOEPALEFFTJDETKSMYFTENMERNCRAPCINOMEDPTKEKNTRFHAINGYI	1064
Db	1759	DSNMPVDMREFFVLTFMFTEDEKSMYVEKKSRSRMR-----LTSGEMKKSHEFHAINGMI	1812
Qy	1065	MDPLPGVVMADOKIRRYTLLSMGSNENHSHIEHSGHVTYAKKKEBYKALNVPQVGET	1124
Db	1813	Y-SLPGIKMTEQEWNVRLHLNLNIGGSODIHVHFHGOITLLENGNQOHQGLVWPLLPESFT	1871
Qy	1125	VEMLPSKAGTWRVYCLLIGENHLHAQMSTLETVYSSKCCQPTLGMASGHIHDPQITASGOYQ	1184
Db	1872	LEKKASKRPGMWMLNTEYGENORACMOPTPELLMDDCRMPMLSTIGLISDSQIKASFEIYG	1933
Qy	1185	WADKLARLHYSGSINAWSTKE--PFS---WIKVDLAPMIIHGICTOGANQKFSLLYIS	1238
Db	1932	WEPLRLALNNGGSYNASVSEKLAIEFASKPMIOYDMQKEVITITGQTOGAKHYLKSCTYT	1991
Qy	1239	QFLIMYSLDGKKMQTYRGNSGTGLMAVFEGNVDSSGICKININPNPIAYIYLHPHYSIR	1296
Db	1992	EFYVAVSSNIMQIEFKGNSSTRNMYFENGNSDASTIKENQDPPVARYIISPRAVNR	2051
Qy	1299	STRJMEMLGCDLNCSPMLGMSKASISDAQTASSYFTNMFAT--WSPSKARLHLOGRNSA	1357
Db	2052	PTLRLELOGCEVNCSTPLGHEKNKRIENKQITASSFKKSMGMDYWEPRALNMOGRYNA	2111
Qy	1358	WRQVNVNPKEMLOVDFOKTKMYKVTGVTTOGVKSLTSMYKEFLISSODGHQWTLFPONG	1417
Db	2112	WQAKNNANNKQMLEIDLTKIKITAIITOGCKSLSEMYKSYTIIHYSBOGVEMKPRILKS	2174
Qy	1418	KV--KVRQGNQDSTPTPVNSLDPELLRYLRIRHQSWHQAIALRNEVLGCEADOLY	1471
Db	2172	SMWDKIEFGNTNTRKGHVKNFNPRIISKFIKRIKRYKTMQSTILRLLEFGC---DIY	2224
RESULT 5			
FA5_BOVIN			
ID	FA5_BOVIN	STANDARD:	PRT: 2211 AA.
AC	Q28107: Q28108:		
DT	01-NOV-1997 (Rel. 35, Created)		
DT	01-NOV-1997 (Rel. 35, Last sequence update)		
DT	01-NOV-1997 (Rel. 35, Last annotation update)		
DE	COAGULATION FACTOR V PRECURSOR (ACTIVATED PROTEIN C COFACTOR).		
GN	F5.		
OS	Bos taurus (Bovine).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;		
OC	Bovidae; Bovine; Bos.		
OX	NCBI_TaxID=9913;		
RN	(1)		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Liver;		
RA	MEDLINE=92147638: Pubmed=1737753:		
RX	Guinco E.R., Esmon C.T., Mann K.G., Macgillivray R.T.:		
RT	"The complete cDNA sequence of bovine coagulation factor V.";		
RL	J. Biol. Chem. 267:2971-2978(1992).		
CC	-1- FUNCTION: COAGULATION FACTOR V IS A COFACTOR THAT PARTICIPATES		
CC	WITH FACTOR XA TO ACTIVATE PROTHROMBIN TO THROMBIN.		
CC	-1- SUBUNIT: FACTOR VA IS COMPOSED OF AN HEAVY CHAIN AND OF A LIGHT		
CC	CHAIN NONCOVALENTLY BOUND. THE INTERACTION BETWEEN THE TWO CHAINS		
CC	IS CALCIUM-DEPENDENT.		
CC	-1- DOMAIN: DOMAIN B CONTAINS 29.5 X 9 AA TANDEM REPEATS, AND 2 X 17		
CC	AA REPEATS.		
CC	-1- PTM: THROMBIN ACTIVATES FACTOR V PROTEOLYTICALLY TO THE ACTIVE		
CC	COFACTOR, FACTOR V(A) (FORMATION OF A HEAVY CHAIN AT THE N-		
CC	TERMINUS AND A LIGHT CHAIN AT THE C-TERMINUS).		
CC	-1- SIMILARITY: CONTAINS 3 F5/8 TYPE A DOMAINS; EACH IS COMPOSED OF		
CC	2 PLASTOCYANIN-LIKE REPEATS.		
CC	-1- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.		

CC	-I- SIMILARITY: STRONG, TO COAGULATION FACTOR VIII.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/or_send_an_email_to_license@isb-sib.ch">http://www.isb-sib.ch/announce/</a>
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).
CC	-----
DR	EMBL; M81440; AAA30512.1; -.
DR	EMBL; M81441; AAA30513.1; -.
DR	HSSP; P00450; IKCW.
DR	InterPro: IPR000421; -.
DR	InterPro: IPR001117; -.
DR	Pfam: PF00394; Cu-oxidase; 3.
DR	Pfam: PF00754; F5_F8_type.C; 2.
DR	PROSITE; PS00079; MULTICOPPER_OXIDASE1; 2.
DR	PROSITE; PS01285; FA58C_1; 2.
DR	PROSITE; PS01286; FA58C_2; 2.
KW	Blood coagulation; Plasma; Glycoprotein; Calcium; Signal; Zymogen;
KW	Repeat.
FT	SIGNAL. 1 28
FT	CHAIN 29 2211
FT	CHAIN 29 741
FT	PEPTIDE 742 1564
FT	
FT	CHAIN 1565 2211
FT	DOMAIN 30 327
FT	DOMAIN 30 193
FT	DOMAIN 203 327
FT	DOMAIN 348 686
FT	DOMAIN 348 525
FT	DOMAIN 535 686
FT	DOMAIN 696 1564
FT	DOMAIN 899 915
FT	SIMILAR 1124 1151
FT	DOMAIN 1124 1137
FT	REPEAT 1138 1151
FT	REPEAT 1188 1453
FT	DOMAIN 1188 1453
FT	
FT	REPEAT 1188 1196
FT	REPEAT 1197 1205
FT	REPEAT 1206 1214
FT	REPEAT 1215 1223
FT	REPEAT 1224 1232
FT	REPEAT 1233 1241
FT	REPEAT 1242 1250
FT	REPEAT 1251 1259
FT	REPEAT 1260 1268
FT	REPEAT 1269 1277
FT	REPEAT 1278 1286
FT	REPEAT 1287 1295
FT	REPEAT 1296 1304
FT	REPEAT 1305 1313
FT	REPEAT 1314 1322
FT	REPEAT 1323 1331
FT	REPEAT 1332 1340
FT	REPEAT 1341 1349
FT	REPEAT 1350 1358
FT	REPEAT 1359 1367
FT	REPEAT 1368 1376
FT	REPEAT 1377 1385
FT	REPEAT 1386 1394
FT	REPEAT 1395 1403
FT	REPEAT 1404 1412
FT	REPEAT 1413 1421
FT	REPEAT 1422 1430
FT	REPEAT 1431 1439
FT	REPEAT 1440 1444
FT	REPEAT 1445 1453
FT	REPEAT 1459 1890
FT	DOMAIN 1569 1738
FT	DOMAIN 1569 1738
FT	
FT	REPEAT 1188 1196
FT	REPEAT 1197 1205
FT	REPEAT 1206 1214
FT	REPEAT 1215 1223
FT	REPEAT 1224 1232
FT	REPEAT 1233 1241
FT	REPEAT 1242 1250
FT	REPEAT 1251 1259
FT	REPEAT 1260 1268
FT	REPEAT 1269 1277
FT	REPEAT 1278 1286
FT	REPEAT 1287 1295
FT	REPEAT 1296 1304
FT	REPEAT 1305 1313
FT	REPEAT 1314 1322
FT	REPEAT 1323 1331
FT	REPEAT 1332 1340
FT	REPEAT 1341 1349
FT	REPEAT 1350 1358
FT	REPEAT 1359 1367
FT	REPEAT 1368 1376
FT	REPEAT 1377 1385
FT	REPEAT 1386 1394
FT	REPEAT 1395 1403
FT	REPEAT 1404 1412
FT	REPEAT 1413 1421
FT	REPEAT 1422 1430
FT	REPEAT 1431 1439
FT	REPEAT 1440 1444
FT	REPEAT 1445 1453
FT	REPEAT 1459 1890
FT	DOMAIN 1569 1738
FT	DOMAIN 1569 1738
FT	
FT	REPEAT 1188 1196
FT	REPEAT 1197 1205
FT	REPEAT 1206 1214
FT	REPEAT 1215 1223
FT	REPEAT 1224 1232
FT	REPEAT 1233 1241
FT	REPEAT 1242 1250
FT	REPEAT 1251 1259
FT	REPEAT 1260 1268
FT	REPEAT 1269 1277
FT	REPEAT 1278 1286
FT	REPEAT 1287 1295
FT	REPEAT 1296 1304
FT	REPEAT 1305 1313
FT	REPEAT 1314 1322
FT	REPEAT 1323 1331
FT	REPEAT 1332 1340
FT	REPEAT 1341 1349
FT	REPEAT 1350 1358
FT	REPEAT 1359 1367
FT	REPEAT 1368 1376
FT	REPEAT 1377 1385
FT	REPEAT 1386 1394
FT	REPEAT 1395 1403
FT	REPEAT 1404 1412
FT	REPEAT 1413 1421
FT	REPEAT 1422 1430
FT	REPEAT 1431 1439
FT	REPEAT 1440 1444
FT	REPEAT 1445 1453
FT	REPEAT 1459 1890
FT	DOMAIN 1569 1738
FT	DOMAIN 1569 1738
FT	
FT	REPEAT 1188 1196
FT	REPEAT 1197 1205
FT	REPEAT 1206 1214
FT	REPEAT 1215 1223
FT	REPEAT 1224 1232
FT	REPEAT 1233 1241
FT	REPEAT 1242 1250
FT	REPEAT 1251 1259
FT	



```

FT DOMAIN 1748 1890 PLASTOCYANIN-LIKE 6.
FT DOMAIN 1894 2048 F5/8 TYPE C 1.
FT DOMAIN 2053 2208 F5/8 TYPE C 2.
FT SITE 741 742 CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
FT SITE 1034 1035 CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
FT SITE 1564 1565 CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
FT DISULFID 167 193 PROBABLE.
FT DISULFID 499 525 PROBABLE.
FT DISULFID 1712 1738 PROBABLE.
FT DISULFID 1894 2048 BY SIMILARITY.
FT DISULFID 2053 2208 BY SIMILARITY.
FT CARBOHYD 225 225 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 239 239 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 297 297 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 382 382 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 460 460 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 553 553 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 587 587 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 745 745 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 756 756 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 774 774 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 780 780 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 902 902 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 952 952 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 964 964 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1044 1044 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1053 1053 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1062 1062 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1071 1071 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1078 1078 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1094 1094 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1451 1451 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1490 1490 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1550 1550 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1690 1690 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1839 1839 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1997 1997 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2196 2196 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT 587 592 NPTLPA -> T (IN VARIANT 2).
SQ SEQUENCE 2211 AA; 248981 MW; CBBF90B738667C45 CRC64;
```

Query Match 30.2%; Score 2379; DB 1; Length 2211;  
Best Local Similarity 26.3%; Pred. No. 1.6e-147;  
Matches 597; Conservative 267; Mismatches 496; Indels 910; Gaps 37;

```

QY 22 RRYVGLAVLSVDYMQSLGELPVDAREPPRYPKSPENTSVYKKTLEFVEFVHLENIA 81
DB 32 RQFYVAASIRWNYR-----PESTHL-----SSKPFETS--FKIYVREYAV-FQKE 76
QY 82 KRPPMGLGPTIOAEYVDIVILKNASHVSLHANGVYWKASGEAEVDDOTSORE 141
DB 77 KFOSTSGLGLEPTLYAEVGDILKVFHKNAHKPLSIHAGIKYKFSSEASYSDDLPMEE 136
QY 142 KEDKVFPGSGHTYVQVYLKENGPMASDPLCLTYLSYSHVDLVKLNLSGILGALLVCREG 201
DB 137 KMDDAVNAAGQETTYEMITSEHSGPETHDPCPLTHIYYSVNLVEDPNSGLIPPLICKKG 196
QY 202 SLAKKKTQTL--HKFILLPAVFDECKSWHSETKNSLMDRODAASARAMEKMTVNGVNR 259
DB 197 TLTEGTOKMEFKQHVLMFAVEDEKSWNQTS-----LMYTVNGYGV 240
QY 260 SLPGGLIGHKRSVYVHYVGMGTPEVHSIFLEGHTFLVRNHRQASLESPITFLAQOLL 319
DB 241 TMDPDTTVAHDIHSHLIGMSGPELFSIHENGQVLEONHRIKISAITTVASVSTTANMTV 300
QY 330 MDLGFLLFCHLISHQHDGMEAYVVKDSCPEEPOLRMKNNEEAEDYDDDLTDESDVYVF 379
DB 301 SPEGWTIASLIPRHFGQGMQAVYIDIKKCAKTRNPK-----LTFMDQ----- 343
QY 360 DDDNSPSTQIRSAKKRPKTVVHTIAAEEDMDYAPLVLAADPDRSYSQYVYLNNGPQIRG 439
DB 344 -----RRHKWEYFIAAEEVIMDYAPIIIPAMDKKYSLSLHIDNSNRIG 388
```

```

QY 440 RKYKKRFMAVYDETEKTR--EAIOHESGLIGLGVGCDTLITFKKQASRPYNIYRH 497
DB 389 KHKKKVYQYODDSFTKREDDSSSGDGLGPIIAQVADTKIVFKNNASYSIYHP 448
QY 498 GIT-----DVRPLYSRRLPKGVKHLKDFPLILGELTFKYMTVTEVGPSPKSPRCL 548
DB 449 GVTFSPYDNEVNSSSTSGSNMTIRAVR-----PGEITYYKWNILEDEDEPENAQCL 500
QY 549 TRYSSFVNMERDLASGLIGPLILCYKESVDQGNQIMSDKRVLLFSVDENRWLYTE 608
DB 501 TRPYYSNVNITRDLASGLIGLILICKRSIDRGRIOAADIEQAAVFAVEDEKSWIYED 560
QY 609 NIQRFPLNAGVQLEDEPEQASNIH-----SINGVFPISQ--LSVCLHEVAWYILISG 662
DB 561 NIYKFCENPEKARKDDPKFYESNIMSNFTLLPAINGVPEISIPILGFCFDYTVQMHFCVSG 620
QY 663 AOTDFLVSFGYTFHKHKKVYEDTLTLFPFSGEYVMSMENPGLMILGCHNSDFRNKMT 722
DB 621 TQNDILITHTFGHSFYGKRHEDTLTLTFMQGESYVTJMDNVGTWMLTTMNSPRSKLR 680
QY 723 ALIKVSCDKNTGD-----YEDS-----YEDISAVL--- 749
DB 681 LRFROAKCIRNDDDDSEYIIEPSCSTAMTKKHDSSEIENDADSDYDELALILGL 740
QY 750 ----- 749
DB 741 RSFRNSLNQEKDELNLTALALEKDEEFLPPSANRLSDNSSRSHVSLIAKNFAESIK 800
QY 750 -----LSKNNAIEP----- 758
DB 801 TLHLLEAPAAGSPLEHAGLDKKSALNPPAHESSPYSEDPREDHPLSDYTVGSLLPFGTG 860
QY 759 -----RSFSQ----- 763
DB 861 FKNRKPAKHQRFQVGRGAQAKHKFSQTRPPAKHTRTLQSODNSSSRMGPMEDIPSDLL 920
QY 764 ----- 763
DB 921 LQOKDPYKILNGBMHLVSEKSEYELIQDANENKYVKLNPSONDSRTNGENIPFNKSHG 980
QY 764 -NSRHPS-----TRQKOFNATP-----PVL 782
DB 981 KQSGHPTFLVTRRRKPLQODNODRRNSLKEGLPLIRTKRKKKEKRPXVHYNPLSPRSHPLR 1040
QY 783 -----KRHOREI-----TRTTLQSDQEEI-----DYDD----- 805
DB 1041 GVNASFSDRRHNHSLILHASNETSLSIDLNOTFPSMNLSLAASLPDHDQTSPTNDTSQT 1100
QY 806 -----TISVEAMKKEDEDDIYDED-----ENQSP----- 827
DB 1101 SSPDPLPYTVSPEEHYQIFPIQDSPTHTSTAPSNKSPDPTHTSTAPSNKSPPTOPSQIP 1160
QY 828 ----- 827
DB 1161 NYDLNRAPIPTVQSIFPSLELVWQTASLDSQPSISPDLGQALMSDPDQGESLSPDL 1220
QY 828 ----- 827
DB 1221 GQTSPLDLSQESLSPDLGQTALSPDPQSQESLSPDLGQTALSPDPQSQESLSPDLGQTALS 1280
QY 828 ----- 827
DB 1281 PDPGQESLSPDLGQTALSPDLDSQESLSPDLGQTALSPDPQSQESLSPDLGQTALSPDPQOE 1340
QY 828 ----- 827
DB 1341 SLSPDLGQTALSPDLGQESLSPDLGQTALSPDPQSQESLSPDLGQTALSPDLGQESLSPDL 1400
QY 828 ----- 827
DB 1401 GQTALSPDLDSQESLSPDLGQTALSPDLDSQESLSPDLGQTALSPDLGQTALSPDLGQTALSPDL 1460
```



```

QY 828 ----- 827
Db 1461 QSLPLPEFGQTFPNADIGQSPSPDSTLNTNFIPEEFNLVYVGLSRDDGYEILIPRQ 1520
QY 828 ----- RSKOKTRHFIAVER 844
Db 1521 KEESSEEDYGEFEFVAYNDPYQDTRTDINSSNPNDAWYLRSNTGNKRYYYIAEEI 1580
QY 845 LMDYSSSPHYLRNAAOGSS---VPO---FKKVFQEOFDGSEFTQPLGENEHLGL 898
Db 1581 SMDYS-----KTVQSDDDVDYVEDIVYKKVYFRKTIYDSTFTKIDPQGEVEHGL 1631
QY 899 GPYIRAEVDNIMVTFERNQASRPYSEYSSLIYE-----BDQROGAPRRNFYKPN 950
Db 1632 GPVIRAEVDVIVQREKKNLASRPYSLAHGLSYEKSEKTYEDDSEPEWKEEDNAIQPNK 1691
QY 951 JTYTYWQVQHHNAPTKDEDFCKAMAFSDVLEKDVHSGILGVLCHTTLNPAHRQY 1010
Db 1692 JTYTYWHAATTRSGPENPGSACRAMAYSAVNPBKDIHSGILGLICRKGTLDKETMPV 1751
QY 1011 TVQEFALFTIFDETKSMYFTEEMERNCAPCNIQMEDPFKENYRPHALNGYIMDTLP 1070
Db 1752 DMREYVLEFMVDEKKSMTYDKPTSMKRASS-----EVKSHERHALNGMAYN-LPG 1804
QY 1071 LVMAODQIRIWMYLLSGNSMENIHSHFSGHVTVRKKEEYKMLYMLPGVFETVEMLP 1130
Db 1805 LRMVQEWVRLHLNLNGSRDIHVHFHGQTLLENGTQHQGLGWPLPLPGSFKTLEKKA 1864
QY 1131 KAGIRVRELLIGENHLAGSTLFVYNSNKCQPLGMAAGHIIROPQTASQYQOMAKLA 1190
Db 1865 KFGWMLDVEGEIQRAGQKOTPLIVDRECKMPMGLSTGLIASQIQOASFQWGWEPKLA 1924
QY 1191 RLHYSGSIAM-----STK-EPFSWIKVDLAPMIIGITIGARQKSFSLYSIQFTIMY 1244
Db 1925 RLNNGSGYAMWIAEKSTEFENPBPWIOVDMQKEVLLTGIOGAKHLKRYTTEFCVAY 1984
QY 1245 SLDGKRWQYRGNSGTGLWVFGNVDSGIGKHNIEPPIIARYIRLPHYTSIRSTLRME 1304
Db 1985 SLDRKMRIRFEKNSSTRNVYFSGNSDASTIKENQIDPPVARYIRISPTSYNKPALRLE 2044
QY 1305 LMGCDLNSCMLPGMESKATISDAQITASSYFTMFPAT-WSPSKARLHLQGRNANWRQVN 1363
Db 2045 LOGCEVNGSTPLGMSGKATENKQITASSFKKSMGWYMEPPLARLMAOGSRVNAWQAKAN 2104
QY 1364 NPEKMLQVDFOKTMKYTGVTQGVSKSLTSMYKEFLISSODGHQWTLTFQNGKV--KV 1421
Db 2105 NNNQWMLQIDLKIKKTAIVTQGCCKLSSEMYKSYTIHSDGDTMCKPYREKSSMWDKI 2164
QY 1422 FQGNDSFTPVVNSLDPPLITRYLRHPOSWVHQIALRMEVLCEADODLY 1471
Db 2165 FEGNNNVNRGHVKNFNPPIITSRIIRIIPKTNOSIALRLLEFGC---DMY 2211

RESULT 6
CERU_HUMAN STANDARD: PRT; 1065 AA.
AC P00450; Q14063;
DT 21-JUL-1986 (Rel. 01, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE CERULOPLASMIN PRECURSOR (EC 1.16.3.1) (FERROXIDASE).
GN Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=86259737; PubMed=2873574;
RA Koschinsky M.L., Funk W.D., Van Oost B.A., McGillivray R.T.A.;
RT "Complete cDNA sequence of human ceruloplasmin.",
RL Proc. Natl. Acad. Sci. U.S.A. 83:5086-5090(1986).
RN [2]

```

```

RP SEQUENCE OF 1-1006 FROM N.A.
RX MEDLINE=95217183; PubMed=7702601;
RA Daimon M., Yamatani K., Igarashi M., Fukase N., Kawanami T.,
RA Kato T., Tomimaga M., Sasaki H.;
RT "Fine structure of human ceruloplasmin gene.",
RL Biochem. Biophys. Res. Commun. 208:1028-1035(1995).
RN [3]
RP SEQUENCE OF 1-40; 549-599; 784-829 AND 919-952 FROM N.A.
RX MEDLINE=86275241; PubMed=3755405;
RA Mercer J.F.B., Grimes A.;
RT "Isolation of a human ceruloplasmin cDNA clone that includes the N-
RT terminal leader sequence.",
RL FEBS Lett. 203:185-190(1986).
RN [4]
RP SEQUENCE OF 218-1065 FROM N.A.
RX MEDLINE=86205876; PubMed=3486416;
RA Yang F., Naylor S.L., Lum J.B., Cusshaw S., McCombs J.L.,
RA Naberhaus K.H., McGill J.R., Adrian G.S., Moore C.M., Barnett D.R.,
RA Bowman B.H.;
RT "Characterization, mapping, and expression of the human ceruloplasmin
RT gene.",
RL Proc. Natl. Acad. Sci. U.S.A. 83:3257-3261(1986).
RN [5]
RP SEQUENCE OF 20-1065.
RX MEDLINE=84119493; PubMed=6582496;
RA Takahashi N., Ortel T.L., Putnam F.W.;
RT "Single-chain structure of human ceruloplasmin: the complete amino
RT acid sequence of the whole molecule.",
RL Proc. Natl. Acad. Sci. U.S.A. 81:390-394(1984).
RN [6]
RP SEQUENCE OF 158-333; 518-724 AND 858-1065.
RX MEDLINE=83117800; PubMed=6571985;
RA Takahashi N., Bauman R.A., Ortel T.L., Dwulet F.E., Wang C.-C.,
RA Putnam F.W.;
RT "Internal triplication in the structure of human ceruloplasmin.",
RL Proc. Natl. Acad. Sci. U.S.A. 80:115-119(1983).
RN [7]
RP SEQUENCE OF 501-905.
RX MEDLINE=81199407; PubMed=6940148;
RA Dwulet F.E., Putnam F.W.;
RT "Complete amino acid sequence of a 50,000-dalton fragment of human
RT ceruloplasmin.",
RL Proc. Natl. Acad. Sci. U.S.A. 78:790-794(1981).
RN [8]
RP SEQUENCE OF 907-1065.
RX MEDLINE=80137543; PubMed=6987229;
RA Kingston I.B., Kingston B.L., Putnam F.W.;
RT "Primary structure of a histidine-rich proteolytic fragment of human
RT ceruloplasmin. I. Amino acid sequence of the cyanogen bromide
RT peptides.",
RL J. Biol. Chem. 255:2878-2885(1980).
RN [9]
RP SEQUENCE OF 907-1065.
RX MEDLINE=80137544; PubMed=6987230;
RA Kingston I.B., Kingston B.L., Putnam F.W.;
RT "Primary structure of a histidine-rich proteolytic fragment of human
RT ceruloplasmin. II. Amino acid sequence of the tryptic peptides.",
RL J. Biol. Chem. 255:2886-2896(1980).
RN [10]
RP SEQUENCE OF 1007-1061 FROM N.A.
RX MEDLINE=90285218; PubMed=2355023;
RA Yang F.M., Friedrichs W.E., Cupples R.L., Banifacio M.J.,
RA Sanford J.A., Horton W.A., Bowman B.H.;
RT "Human ceruloplasmin. Tissue-specific expression of transcripts
RT produced by alternative splicing.",
RL J. Biol. Chem. 265:10780-10785(1990).
RN [11]
RP X-RAY CRYSTALLOGRAPHY (3.1 ANGSTROMS).
RA Zaitseva I., Zaitsev V., Card G., Moshkov K., Bax B., Ralph A.,
RA Lindley P.;
RT "The X-ray structure of human serum ceruloplasmin at 3.1 A: nature of
RT the copper centres.",
RL J. Biol. Inorg. Chem. 1:15-23(1996).

```



[illegible]







FT	CARBOHYD	756	756	N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT	CARBOHYD	920	920	N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT	CONFLICT	271	271	ED -> A (IN REF. 2).
FT	CONFLICT	604	605	GD -> DN (IN REF. 2).
FT	CONFLICT	823	823	T -> S (IN REF. 2).
FT	CONFLICT	833	833	V -> L (IN REF. 2).
FT	CONFLICT	868	868	C -> V (IN REF. 2).
FT	CONFLICT	868	891	L -> R (IN REF. 2).
FO	SEQUENCE	1059	AA; 120840	MM; 12BA3B990A0AB95E3 CRC64;

Query Match	21.3%	Score 1678	DB 1:	Length 1059
Best Local Similarity	32.0%	Pred. No. 4.6e-102		
Matches	385	Conservative 188	Mismatches 429	Indels 200
				Gaps 23

```

QY 5 LSTCFCLLPFCESATRRYYLGAVELSMDYMO.-SDLCEL.-PVDARFPREVRKSEPFNTS 62
Db 6 L$ALLFTL-H$SIAMT$REKHYYIGLIEAVMDYASGSEEEKLISVDTPEQ$NFYLRNGPDRIG 64
QY 63 VVYKTLTVE$T$VHL$FNIAKPRPPM$GILLPTIOAEVVDYVITLKNASHV$SLHAYV 122
Db 65 RRYK$ALTS$EYTDGT$T$T$ITDKPAVL$GEL$GVP$IKAEV$D$K$SV$H$K$N$K$ASRP$YTH$AGV 124
QY 123 SYMK$SEGA$EYDDO$T$Q$REKED$K$F$P$G$S$H$Y$Q$V$Y$K$EN$G$P$MA$SDP$CLT$Y$SL$SHD 182
Db 125 TYT$AN$E$A$II$P$ND$T$D$Q$R$AD$K$F$P$Q$O$Y$Y$V$L$R$A.-NE$P$G$G$D$S$C$V$RI$TH$SHD 183
QY 183 LVKD$N$SCL$IC$ALL$V$C$R$E$G$SL$A$K$E$K$T$O$T$L.-H$K$ILL$P$AV$F$D$E$K$S$H$E$T$K$N$SLM$---- 236
Db 184 APK$IAS$OL$IP$ILL$C$K$K$G$S$L$H$K$E$E$EN$ID$O$F$V$M$F$S$V$E$N$L$S$M$Y$LED$N$IK$T$P$C$SE$P 243
QY 237 -QDD$A$S$AR$AP$M$K$H$T$NG$V$Y$N$R$SL$P$LL$IG$H$R$K$S$Y$M$H$Y$G$M$ST$E$V$H$S$IF$LEG$H$F 295
Db 244 KYVD$K$D$NE$F$Q$S$N$K$Y$S$ING$Y$T$F$G$SL$P$G$LS$M$CA$E$D$R$V$K$Y$M$Y$LG$M$K$E$N$V$D$H$E$L$F$H$G$AL 303
QY 296 LV$R$H$R$Q$A$S$LE$I$P$T$F$T$LA$Q$T$L$M$D$IO$F$L$E$F$CH$ISS$H$O$H$D$M$A$Y$K$V$D$S$C$P$E$P$Q$R 355
Db 304 T$K$N$Y$H$T$DI$M$L$F$AT$IL$D$V$S$V$AQ$N$P$G$V$M$M$L$S$C$N$L$H$K$AG$IA$F$Q$V$R$D$C$----- 356
QY 356 M$K$N$NE$A$D$Y$D$D$D$L$T$S$M$D$Y$V$R$P$D$D$N$S$P$S$T$O$I$R$S$A$K$K$P$K$W$H$Y$IA$E$E$D$M$Y$A 415
Db 357 ---NK$P$P$-D$D$D$I$O$D$R$H$V$-----R$H$-----Y$Y$IA$E$E$L$T$W$D$A 385
QY 416 P-----LV$L$AP$D$R$S$Y$K$S$O$Y$LN$N$G$P$O$R$IG$K$K$Y$V$F$M$A$Y$T$P$E$T$---K$T$R$E$A$ 461
Db 386 P$G$T$D$T$F$G$E$N$F$T$S$G$S$R$V$F$E$D$---G$A$T$R$IG$S$T$K$K$LY$R$E$Y$T$D$S$F$N$K$R$E$G$D 441
QY 462 Q$H$E$S$IG$LP$LL$Y$G$E$V$D$T$LL$I$F$K$N$Q$A$S$P$R$Y$N$H$G$IT$D$V$R$---P$Y$S$R$LL$P$K$Y$K$H$L 517
Db 442 E$H$H$G$IL$C$P$V$IA$E$G$D$II$R$V$T$F$H$N$K$Q$F$P$LS$I$Q$M$G$V$R$F$T$K$E$N$G$Y$Y$G$---P$D$G$R$S$K 498
QY 518 K$D$P$IL$P$E$I$K$K$W$Y$V$E$D$G$P$T$K$S$D$R$C$IL$T$R$Y$S$S$V$M$N$E$R$D$LA$S$IG$LL$C$Y$K$S 577
Db 499 O$A$S$H$A$P$E$T$Y$T$M$Y$T$P$K$E$M$G$P$Y$A$D$V$C$LS$K$M$Y$S$E$V$D$L$K$D$LT$G$IL$G$M$K$IC$K$G$S 558
QY 578 V$D$Q$G$N$Q$I$M$S$D$R$N$V$IL$F$S$V$E$D$E$N$S$W$Y$L$E$N$I$O$F$LE$P$N$P$A$G$V$O$L$E$D$E$F$Q$A$N$I$M$H$S$IN 637
Db 559 L$AD$Q$K$O$D$V$O$K$E$Y$L$F$A$T$V$E$D$E$N$S$ILL$D$N$IR$M$F$T$A$P$E$N$V$D$K$E$D$D$E$D$Q$S$N$K$H$M$SN 618
QY 638 G$Y$V$D$S$I$Q.-L$S$V$C$H$E$V$A$Y$W$IT$LS$GA$O$T$D$E$L$S$V$F$S$G$Y$T$E$K$H$K$W$Y$E$D$T$LL$P$P$S$G$E$T 696
Db 619 G$F$M$G$N$P$E$G$LM$C$G$E$S$I$W$Y$L$F$S$A$G$N$A$D$V$H$G$I$F$S$G$N$Y$L$S$K$E$R$D$T$A$L$F$P$H$K$S$LT 678
QY 697 V$F$M$E$M$E$N$G$L$M$IL$C$H$N$S$D$E$F$R$N$G$R$T$A$L$K$Y$S$C$O$K$N$G$D$Y$E$D$Y$E$D$IS$A$V$L$E$K$N$N$AI 756
Db 679 L$M$T$P$D$E$G$S$D$V$E$C$IL$T$D$H$T$G$K$K$O$K$Y$Y$N$Q$C$-----K$O$F$E$D$Y$T$L$----- 722
QY 757 E$P$R$F$S$O$N$S$R$H$P$T$R$O$K$O$F$A$N$T$P$V$L$K$H$Q$E$R$IT$F$T$LL$O$S$Q$E$E$ID$Y$D$D$T$J$S$E$N$K$K$E$D$F 816
Db 723 ----- 722
QY 817 D$Y$E$D$E$N$Q$S$R$F$O$K$T$R$H$F$IA$V$E$L$M$D$Y$G$M$S$S$H$V$L$R$N$K$O$S$S$Y$----- 866

```

```
Db      723 -----QGERTYYIAAVEVEMDYSPPSRDMEMELAHNIOEUVNSNAFLDKREE 767
Qy      867 ---FOFKVVYOEFPDGSFTQPLRYGELNEHGLIGPTRYAEVDINIMTFRNQASRPS 923
       :|::||::||::||::||::||::||::||::||::||::||::||::||::||::||:
Db      768 FIGSKYKKVVYREFDSFFREOVKRRAREEHGMLGPLIHADVGAKVKVFKNMATRPS 827
Qy      924 FYSSLISFEEDQRGAEBRKNNPKV---NENFTYFMKVQHNAHPRKDEDDCKAMAFESDY 980
       ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      828 IHA-----HGVKTKSSVTAPDLGEVRTYIMQIPERSAGTSDSPCPMAVYSTV 877
Qy      981 DLEKDVHGSLIGPLLVC---HTNTLPANAGROVYTFEPALFETIDETFSWTFTEMMEN 1037
       |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      878 DRVAVLYSGLLGPRLCYCKSKSYKVNPK-----KKMESLFLVEDENSWLDDININTY 932
Qy      1038 CRAPCNLOMEDPTPEKENYRFPAINGIYMDLTPLGYLAADQIRIWLXLSGNSENHSIH 1097
       |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      933 PDHEEKDNKDNDDEEFESSNKMHAINKMGNLOGLTIMHGDENVYVMAGNGEIDLTHVHF 992
Qy      1098 SGHFTYVKKKEEYKALNLNPGVEFYEMLRPSKRGIRVELLIEGHLAAGSTLFVYS 1157
       ||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      993 HGHSEFYQNHRISSDVPDFPFPGYTLEMEFGTPGTWLLHCYVDHIHAGVTTYTVLP 1052
Qy      1158 NK 1159
       !:
Db      1053 NQ 1054

RESULT      8
CERU MOUSE
AC          STANDARD;      PRT: 1062 AA.
ID          061147;
DT          01-NOV-1997 (Rel. 35, Created)
DT          01-NOV-1997 (Rel. 35, Last sequence update)
DT          15-JUL-1999 (Rel. 38, last annotation update)
DE          CERULOPLASMIN PRECURSOR (EC 1.16.3.1) (FERROXIDASE).
GN          CP.
OS          Mus musculus (Mouse).
OC          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN          NCBI_Taxid=10090;
RP          [1]
RA          Ktomp L.W.J., Farhangrazi Z.S., Choi D.W., Gitlin J.D.;
RL          Submitted (MAR-1996) to the EMBL/Genbank/DBJ databases.
RN          [2]
TT          TISSUE SPECIFICITY.
RA          Ktomp L.W.J., Farhangrazi Z.S., Dugan L.L., Gitlin J.D.:
RT          "Ceruloplasmin gene expression in the murine central nervous system.";
RL          J. Clin. Invest. 98:207-215(1996).
CC          - FUNCTION: CERULOPLASMIN IS A BLUE, COPPER-BINDING (6-7 ATOMS PER MOLECULE) GLYCOPROTEIN FOUND IN PLASMA. FOUR POSSIBLE FUNCTIONS ARE FERROXIDASE ACTIVITY, AMINE OXIDASE ACTIVITY, COPPER TRANSPORT AND HOMEOSTASIS, AND SUPEROXIDE DISMUTASE ACTIVITY.
CC          - CATALYTIC ACTIVITY: 4 FE(2+) + 4 H(+) + O(2) = 4 FE(3+) + 2 H(2)O.
CC          - COFACTOR: BINDS 6 CU-IONS PER MOLECULE. THIS PROTEIN BELONGS TO THE MULTICOPPER OXIDASES WHICH CONTAIN THREE DISTINCT CU CENTERS KNOWN AS TYPE 1 OR BLUE, TYPE 2 OR NORMAL, AND TYPE 3 OR COUPLED BINUCLEAR.
CC          - TISSUE SPECIFICITY: MANY TISSUES, INCLUDING LIVER, EYE AND BRAIN.
CC          - SIMILARITY: CONTAINS 3 FS/8 TYPE A DOMAINS; EACH IS COMPOSED OF 2 PLASTOCYANIN-LIKE REPEATS.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on lists use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
```



DR MGD; MGI:88476; Cp.  
 DR InterPro; IPR001117; -  
 DR InterPro; IPR002355; -  
 DR Pfam; PF00394; Cu-oxidase; 3.  
 DR PROSITE; PS00079; MULTICOPPER\_OXIDASE1; 3.  
 DR PROSITE; PS00080; MULTICOPPER\_OXIDASE2; 1.  
 DR Oxidoreductase; Copper; Metal-binding; Glycoprotein; Plasma; Repeat;  
 KW Signal.  
 FT SIGNAL 1 19  
 FT CHAIN 20 1062  
 FT DOMAIN 20 356  
 FT DOMAIN 20 189  
 FT DOMAIN 208 356  
 FT DOMAIN 369 713  
 FT DOMAIN 369 713  
 FT DOMAIN 565 713  
 FT DOMAIN 725 1057  
 FT DOMAIN 725 896  
 FT DOMAIN 904 1057  
 FT DISULFID 173 199  
 FT DISULFID 275 356  
 FT DISULFID 529 555  
 FT DISULFID 632 713  
 FT DISULFID 870 896  
 FT METAL 120 120  
 FT METAL 122 122  
 FT METAL 179 179  
 FT METAL 181 181  
 FT METAL 980 980  
 FT METAL 993 993  
 FT METAL 995 995  
 FT METAL 1035 1035  
 FT METAL 1036 1036  
 FT METAL 1037 1037  
 FT METAL 1041 1041  
 FT METAL 1046 1046  
 FT CARBOHYD 138 138  
 FT CARBOHYD 226 226  
 FT CARBOHYD 583 583  
 FT CARBOHYD 625 625  
 FT CARBOHYD 757 757  
 FT CARBOHYD 922 922  
 SO SEQUENCE 1062 AA; 121159 MW; F3F52ED09A238F16 CRC64;  
 Query Match 20.8%; Score 1639; DB 1; Length 1062;  
 Best Local Similarity 31.7%; Pred. No. 1.6e-99;  
 Matches 383; Conservative 188; Mismatches 437; Indels 200; Gaps 24;  
 QY 10 FLCLRFPC-----SATRRYLAGVELSWDMQ--SDLGELPVDARPPRPVKSPFN 60  
 DB 3 FLLESTFIFLYSSLALARKHYFICITAEVMDASGTEKKLISVDTEQSNFLQNGPDR 62  
 QY 61 TSVVKKTLFVEFTVHLLENIAKRPWMGLIPTIOAEVDTVAITLLKNASHPVSLAHV 120  
 DB 63 IGRKTKKALFYFEYTDGTEFTSKTDKIPAWLGLGVIAAEVDKYYVHLKKNLASIYTFHAH 122  
 QY 121 GVSVMKASGEAEYDQTSOREKEDKVPFGSHYVQVLEKNGPMASDPLCLTYSLSH 180  
 DB 123 GYTYKKEYGAYYPRDNTTFQRRADKVLPGQOYVYLHA-NEPSPREGDSNCTRIYHSH 181  
 QY 181 VDLVKDNLGSLGALLVCREGSLAKERTQL-HKFTLLFAVFDEGKSWHSETKNSLM-- 236  
 DB 182 VDAPKDIASGLIGPLILCKKSLYKEKKNIDQEFVLFMSVYDENLSWYLEDNIKTFCS 241  
 QY 237 ---QORDASARAPKMHVNGVNSLPGLICHRKSYVWHYIGCTTPEVHSIFLEGH 293  
 DB 242 PEKVYKDEDFQESNMYSINGTTFGSLPGLSCADRVKMYLFGNGENEDVHSAFPHGQ 301  
 QY 294 TFLVNRHROASLEISITFLTAQTLLMDLGOFLPLFCHISHQHDGAEVAYKVVSCEEPQ 353  
 DB 302 ALTSKNYQTDITINLPATLIDAYMAQNPGVWMLSCONLHLKAGLQAFQVODC----- 356

QY 354 LRKKNNEAEDYDDDLTDEMDYVREDDNSPSFQIRSVAKKHPTWVHYTAAEEDWD 413  
 DB 357 -----NKSSKD-----NIRGKHVH-----YITAAEEVIN 383  
 QY 414 YAP-LVLAADDR-----SYKSOYLANNGPORIGRKKKKRPMAYTDETF-----KTRALIGH 463  
 DB 384 YAPSGIDIFTEEEKLTLASGSDSGVFEDGARRIGSYKKMAVREYTGSTNRKREKPPDE 443  
 QY 464 ESGILGLTGEAGDPTLLIFKKNQASRPYNIYPHGIT-----DVRPLYSRRLPGVYHLKD 519  
 DB 444 HLGILGLVIAAEVGDITIKYTFHNKGQHLHSLQMGVSPFAENNGITYG---EPGASSQQA 500  
 QY 520 PP-ILPGEIFKRYKWTYVEDGPTRCDPLTRYYSFVNNERDLASGLIGPLILCYKESV 578  
 DB 501 ASHAPKXXTYETWYVPEKMGPTIYADPVCLSKMYSAVPTKDIIFGLIGPMKICKKSL 560  
 QY 579 DQGNQIMSKRNVILFSPVDENRSWTLFPIENQRLPNAGVQLDEPPEQASINIMHSING 638  
 DB 561 LADGRQDVDKERYLEFPTVDENESLILDDINIMFTHADPQVDKEDQESNKKHSMNG 620  
 QY 639 YVFDLSQL-SVCLHEVAYWYILSIGAQDPLSVFSGYTFKHKMYVEDTLTFPFSGETV 697  
 DB 621 FMYGNOGMPHMCLEGSIVWYLFSGAGNADVHGIFSGNTYLCGGERDTANLFPKHSITL 680  
 QY 698 FMSMENGLMIILCCHNSDFRNKMTALLKVSQCKNTGDIYEDSYEDISAYLLSKNALE 757  
 DB 681 LMPDFTGTEVDCLTTHDYTGKMKQKYVNVQ----- 713  
 QY 758 PRFSQNSRHPSTRQOFANATPVYLRHQREITRTLOSQDEIDYDITISVEMKKEPD 817  
 DB 714 -----ORQF----- 721  
 QY 818 IYDEDENQSPRSFOKTRHYFIAVERLMDYGMSSPHYLNRNAGSV-----EDPT 721  
 DB 722 VYIGE-----RTYVDAVEVMDSPSRAMEKEDLHLQDQNSVNFLLDKEEF 769  
 QY 867 --DQKRVVQFEPTDGSFTQPL-YRGELEHGLGLPYIRAEVEDINWTFRNQASRPYS 923  
 DB 770 IGSYKRVVYQRTDSSFRQVYKRAEEDBHGLIIPPIHANVGDVKVYFKMATRPYS 829  
 QY 924 FYSLSLSYEDQAGAPRKNFY---KPNETKYFFKQVGHNAKPKDEDFCKAMAYFSV 980  
 DB 830 IHA-----HGVTSSIVPTLPGEVATYVQIERSGAGEDSACIPMAVYSIV 879  
 QY 981 DLEKDVHSLIGPLILVC---HTNLTNPARGROYTVOEALFTFIDETSKSVFTENMERN 1037  
 DB 880 DRYKDLISGLIGPLIYCRKSYKAVFSPK-----KKNFEFLFLYVDEPNSWYLDINIKY 934  
 QY 1038 CRAPCNIQMEDPTFKENYRPHAINGYIMDTLPGLVAAQDQRIKWTLLSGNSMENHSIH 1097  
 DB 935 SEHPREKYNKDNBEFLSNKMHAINGMFGNLQLTGMHVDEVMWYLMGNGNEIDLTWHF 994  
 QY 1098 SGHVFYRKKEEYKMLYNLYPGVEFTEVEMLPKAGIMWVECLIGLHLAGSTLFIYS 1157  
 DB 995 HGHSFQYKRGVYSSDVFLEPCTYOTLEKFPOTPGTWHLCHYVDHVAAGATTYVLP 1054  
 QY 1158 NKCOTPLG 1165  
 DB 1055 VEGETKSG 1062  
 RESULT 9  
 MFGM\_RAT  
 ID MFGM\_RAT STANDARD; PRT: 427 AA.  
 AC P70450;  
 DT 01-NOV-1997 (rel. 35, Created)  
 DT 01-NOV-1997 (rel. 35, Last sequence update)  
 DT 30-MAY-2000 (rel. 39, Last annotation update)  
 DE LACTADHERIN PRECURSOR (MILK FAT GLOBULE-BGF FACTOR 8) (MFG-B8) (O-  
 ACETYL GD3 GANGLIOSIDE SYNTHASE) (AGS) (MFGM).  
 GN MFG8 OR AGS.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;







DR Pfam: PF00754; F5\_F8\_type\_C; 2.  
DR PRINTS: PR00010; EGFBL00.  
DR PROSITE: PS00022; EGF\_1; 2.  
DR PROSITE: PS01186; EGF\_2; 2.  
DR PROSITE: PS01285; FA58C\_1; 2.  
DR PROSITE: PS01286; FA58C\_2; 2.  
KW Signal; Glycoprotein; Repeat; EGF-like domain; Milk.  
FT SIGNAL 1 22  
FT CHAIN 23 463 LACTADHERIN.  
FT DOMAIN 24 61 EGF-LIKE 1.  
FT DOMAIN 64 108 EGF-LIKE 2.  
FT DOMAIN 148 303 F5/8 TYPE C 1.  
FT DOMAIN 308 463 F5/8 TYPE C 2.  
FT SITE 87 89 CELL ATTACHMENT SITE (POTENTIAL).  
FT DISULFID 28 39 BY SIMILARITY.  
FT DISULFID 33 49 BY SIMILARITY.  
FT DISULFID 51 60 BY SIMILARITY.  
FT DISULFID 68 79 BY SIMILARITY.  
FT DISULFID 73 96 BY SIMILARITY.  
FT DISULFID 98 107 BY SIMILARITY.  
FT DISULFID 148 303 BY SIMILARITY.  
FT DISULFID 290 294 BY SIMILARITY.  
FT CARBOHYD 308 463 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 61 61 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 266 266 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 316 316 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 426 426 S -> F (IN REF. 2).  
FT CONFLICT 30 30 S -> D (IN REF. 1; AA SEQUENCE).  
FT CONFLICT 35 35 EFNYYNLDGEVFTTAVPTAVPTAPPAFPDLSNNLASR ->  
FT CONFLICT 110 147 G (IN REF. 2).  
FT CONFLICT 168 168 Y -> S (IN REF. 2).  
FT CONFLICT 196 196 L -> T (IN REF. 2).  
FT CONFLICT 309 309 L -> S (IN REF. 2).  
FT CONFLICT 395 395 E -> A (IN REF. 2).  
SQ SEQUENCE 463 AA; 51465 MW; D7B86CEFBAT724D CRC64;  
  
Query Match 8.3%; Score 657; DB 1; Length 463;  
Best Local Similarity 42.2%; Pred. No. 1e-35;  
Matches 135; Conservative 63; Mismatches 110; Indels 12; Gaps 5;  
  
QY 1157 SNKQTPGMAHGHRDPQTASGQY-----GQMAPKLARLHYSSTINAW--STKEPS 1208  
DB 145 ASRCSTQJMGEGALDSDISASVYMGFMGLQRMGPRLARLYRTSIVAMHANSNDKRP 204  
QY 1209 WIKVDLAPMIHIGIKTQGARQKFSLSYISOFIIMYSLDGKMQTFRGNSTGLMVFQGN 1268  
DB 205 WIQVNLKMRVSGVMTQGSASRAGRAEYLKTFKVAISLDGRKEFIQDESGGD-KEFLGN 263  
QY 1269 VDSGIRKINIEPPIIARYIRLPHYSIRSLRMELMGCDLNSCMLPGMESKAISDAQ 1328  
DB 264 LDNNSLTKVMFPTLEAGYIRLYPVSCHRGCTLRFELLGCELLHGCLEPIGLKNNKNTIPDSQ 323  
QY 1329 ITASSYFT--NMFA-TWSSKARLHLQGRSNMARPQVNPKEQLVDPOKMTVGTQV 1385  
DB 324 MSASSSYTWNIRAFGWPHLGRLDNQGKINAMTAQNSAKMELQVLDLQRTQVYTIITQ 383  
QY 1386 GVKSLITSMYEFLLISSQDGHQWTLFQNGKVKVFGQGNQDSFTPVVNSLDPLLTRYL 1445  
DB 384 GARDGHHQYVESYKVAISDDQVMTVYEQSSKVFQGNLDNNSHKHNIIEKPFMARVY 443  
QY 1446 RIHPQSWHQAIRMEVLGC 1465  
DB 444 RVLPSVSHNRITRLRELLGC 463  
  
RESULT 11  
MEGM\_PIG STANDARD: PRT; 409 AA.  
AC P79385;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE LACTADHERIN (MILK FAT GLOBULE-EGF FACTOR 8) (MEG-E8) (MEGM) (SPERM  
GN SURFACE PROTEIN SP47) (PP47).  
GN MEG8.  
OS Sus scrofa (Pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.  
OX NCBI\_TaxID=9823;  
RN (1)  
RP SEQUENCE FROM N.A.  
RC TISSUE-Tests;  
RA Ensslin M.A.;;  
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: MAY BE INVOLVED IN PHOSPHOLIPID BINDING. ZONA PELLUCIDA-  
CC BINDING PROTEIN.  
CC -1- SUBCELLULAR LOCATION: PERIPHERAL MEMBRANE PROTEIN.  
CC -1- TISSUE SPECIFICITY: MAMMARY EPITHELIAL CELL SURFACES AND  
CC SPERMATOZOON.  
CC -1- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.  
CC -1- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb.ch/announce/>  
CC or send an email to [license@isb.sib.ch](mailto:license@isb.sib.ch)).  
CC -----  
DR EMBL: Y11683; CAA72379.1; -  
DR HSSP: P00740; 11XA.  
DR InterPro: IPR000421; -  
DR InterPro: IPR000561; -  
DR Pfam: PF00008; EGF; 2.  
DR Pfam: PF00754; F5\_F8\_type\_C; 2.  
DR PROSITE: PS00022; EGF\_1; 2.  
DR PROSITE: PS01186; EGF\_2; 2.  
DR PROSITE: PS01285; FA58C\_1; 2.  
DR PROSITE: PS01286; FA58C\_2; 2.  
KW Glycoprotein; Repeat; EGF-like domain.  
FT DOMAIN 2 41 EGF-LIKE 1.  
FT DOMAIN 44 88 EGF-LIKE 2.  
FT DOMAIN 91 247 F5/8 TYPE C.  
FT DOMAIN 252 409 F5/8 TYPE C.  
FT SITE 67 69 CELL ATTACHMENT SITE (POTENTIAL).  
FT DISULFID 6 17 BY SIMILARITY.  
FT DISULFID 11 29 BY SIMILARITY.  
FT DISULFID 31 40 BY SIMILARITY.  
FT DISULFID 91 247 BY SIMILARITY.  
FT DISULFID 234 238 BY SIMILARITY.  
FT DISULFID 252 409 BY SIMILARITY.  
FT CARBOHYD 41 41 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 372 372 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 409 AA; 45725 MW; BOC07AF80029927A CRC64;  
  
Query Match 8.3%; Score 650; DB 1; Length 409;  
Best Local Similarity 39.3%; Pred. No. 2.5e-35;  
Matches 139; Conservative 62; Mismatches 125; Indels 28; Gaps 6;  
  
QY 1138 ECLIGHLAGSTLFLVYSNK-----CQPLPGMAHGHRDPQTASGQY-- 1182  
DB 58 ECEVIDDAHRG--DVETETICCKPHQYTGICEIILNAPLMETGAIDAFQISASSMHLG 115  
QY 1183 ----GQMAPKLARLHYSSTINAW--STKEPSWIKVDLAPMIHIGIKTQGARQKFSLSY 1236  
DB 116 FMGLQRMAPELRLHRAGLVNAWMTASNDRNPAWIOVNLRRRVAGVYTGQSRGSAEY 175  
QY 1237 ISOFTIMYSLDGKMQTFRGNSTGLMVFQGNVDSGIRKINIEPPIIARYIRLPHYS 1296  
DB 176 MKTFKVAYSTDRKRFQFIGAESGDKIFPMGLNDSGLKVLNLFEPVLEQVYRLVPIICH 235  
QY 1297 INSTLRMELMGCDLNSCMLPGMESKAISDAQITASSYFTN--MEATWSPSKARLHLQ 1353



```

DB 236 RCGTTLFELLGELGELGCEPLGLKNDNTIPNKQITASSFRTWGLSAFWSWPFYARLDNG 295
OY 1334 RSNARPOVNNKEMLOVDFOKTMVGTGQVKSILTSMYKFEPLISSODGQWLF 1413
DB 296 RFNATTAOSNSSEMLQIDLSGSRVATGILITGGARDFGHIOVAAVKAAYSDDGVSWEX 355
OY 1414 FONGKV--KVFQNDSPFPVNSLDPPLTFRYLRIHPOSWVHQAIALMEVLGC 1465
DB 356 RROGALBEKIFPGNLDNNSHKMFEPLTFRVILPEVAMHNRITLVELLGC 409

RESULT 12
MFGM_BOVIN STANDARD: PRF: 427 AA.
AC 095114: Q27959; P79344;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE LACTADHERIN PRECURSOR (MILK FAT GLOBULE-EGF FACTOR 8) (MFG-E8)
DE (MG57/53) (PAS-6/PAS-7 GLYCOPROTEIN) (MFGM) (SPERM SURFACE PROTEIN
DE SP47) (BP47) (COMPONENTS 15/16).
CN MFGB8.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A., PARTIAL SEQUENCE, AND CARBOHYDRATE-LINKAGE SITES.
RC STRAIN=HOLSTEIN; TISSUE=Mammary gland;
RX MEDLINE=97008954; PubMed=8856064;
RA Hvarregaard J., Andersen M.H., Berglund L., Rasmussen J.T.,
RA Petersen T.E.;
RA "Characterization of glycoprotein PAS-6/7 from membranes of bovine
RT milk fat globules.";
RT Eur. J. Biochem. 240:628-636(1996).
RN [2]
RP SEQUENCE OF 18-427 FROM N.A.
RC TISSUE=Mammary gland;
RX MEDLINE=96125736; PubMed=8541316;
RA Aoki N., Kishi M., Taniguchi Y., Adachi T., Nakamura R.,
RA Matsuda T.;
RA "Molecular cloning of glycoprotein antigens MG57/53 recognized by
RT monoclonal antibodies raised against bovine milk fat globule
RT membrane.";
RT Biochim. Biophys. Acta 1245:385-391(1995).
RN [3]
RP SEQUENCE OF 19-427 FROM N.A.
RC TISSUE=Testis;
RA Ensslin M.A.;
RA Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 140-146; 174-187; 233-246 AND 422-427.
RC TISSUE=Milk;
RX MEDLINE=93250576; PubMed=8485470;
RA Mather I.H., Banghart L.R., Lane W.S.;
RA "The major fat-globule membrane proteins, bovine components 15/16 and
RT guinea pig GP 55, are homologous to MFG-E8, a murine glycoprotein
RT containing epidermal growth factor-like and factor V/VIII-like
RT sequences.";
RL Biochem. Mol. Biol. Int. 29:545-554(1993).
CC -1- FUNCTION: PROBABLY ASSOCIATED WITH PHOSPHOLIPIDS ON THE SURFACE OF
CC MAMMARY EPITHELIAL CELLS AND MILK FAT GLOBULES. ZONA PELLUCIDA-
CC BINDING PROTEIN.
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
CC SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING. THE SHORT FORM
CC LACKS 53 AMINO ACIDS WITHIN THE F5/8 TYPE C 1 DOMAIN.
CC -1- TISSUE SPECIFICITY: MILK AND SPERMATOZOAN.
CC -1- PM: THE 2 O-LINKED GLYCANS CONSIST OF GAL, GLCNAC AND FUC, WITH
CC PROBABLY FUC AS REDUCING TERMINAL SUGAR.
CC -1- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.

```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb.ch).
CC -----
DR EMBL; X91895; CA62997.1; -
DR EMBL; S80643; AAB35894.2; -
DR EMBL; Y11719; CA72406.1; -
DR HSSP; P00740; IIXA.
DR InterPro; IPR000421; -
DR InterPro; IPR000561; -
DR Pfam; PF00008; EGF_2.
DR Pfam; PF00754; F5_F8_type_C; 2.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 2.
DR PROSITE; PS01285; FA58C_1; 2.
DR PROSITE; PS01286; FA58C_2; 2.
KW Signal; Glycoprotein; Milk; Repeat; EGF-like domain;
KW Alternative splicing.
FT SIGNAL 1
FT CHAIN 18
FT DOMAIN 19 427 LACTADHERIN.
FT DOMAIN 20 59 EGF-like 1.
FT DOMAIN 62 106 EGF-like 2.
FT DOMAIN 109 265 F5/8 type C 1.
FT DOMAIN 270 427 F5/8 type C 2.
FT SITE 85 87 CELL ATTACHMENT SITE (POTENTIAL).
FT DISULFID 24 35 BY SIMILARITY.
FT DISULFID 29 47 BY SIMILARITY.
FT DISULFID 49 58 BY SIMILARITY.
FT DISULFID 66 77 BY SIMILARITY.
FT DISULFID 71 94 BY SIMILARITY.
FT DISULFID 96 105 BY SIMILARITY.
FT DISULFID 109 265 BY SIMILARITY.
FT DISULFID 252 256 BY SIMILARITY.
FT DISULFID 270 427 BY SIMILARITY.
FT CARBOHYD 27 27 O-LINKED (FUC. . .) (IN PAS-6).
FT CARBOHYD 34 34 O-LINKED (FUC. . .) (IN PAS-7).
FT CARBOHYD 59 59 N-LINKED (GLCNAC. . .) (HYBRID) (IN PAS-6
FT CARBOHYD 59 59 AND PAS-7).
FT CARBOHYD 227 227 N-LINKED (GLCNAC. . .) (HIGH MANNOSE) (IN
FT CARBOHYD 227 227 PAS-6).
FT VARSPIC 169 221 MISSING (IN SHORT ISOFORM).
FT CONFLICT 19 19 A -> F (IN REF. 1).
FT CONFLICT 28 28 L -> Q (IN REF. 1).
SQ SEQUENCE 427 AA; 47411 MW; 4CBBE3A1DC4EB24 CRC64;

Query Match 8.1%; Score 635; DB 1; Length 427;
Best Local Similarity 37.9%; Pred. No. 2.5e-34;
Matches 135; Conservative 66; Mismatches 123; Indels 32; Gaps 7;

OY 1138 RCLGHEHLHAGKSTFLFLYNNK-----CQPLGASGHIRDPOTASGOY-- 1182
DB 76 ECQVTDSDHRG--DFVIOYICKPLGIYGHICETTCSPLEMQGALADSOISASSMHLG 133
OY 1183 -----GQAPKLARLHYSGSINAWST---KEPFWIKYDLAPLPIHGIKTGAROKFS 1234
DB 134 FMGLQRMWAPELARLHQGTIVNAWTSNGVNDKP--WIGVNLKRMKVVGVTVQASRAGSA 191
OY 1235 LYISQFTIMSLDGKKWQTYGNSGTGLMAFFGVNDSSGICAHNINFPPIIARIYILHPT 1294
DB 192 EYLKTFKVAYSTDFGQFOFQIVAGRSQDKIFITGVNNSGLKINLFDPLETOYVALVPII 251
OY 1295 YSINSTLMELMGCLNCSMPLGMSKAISDAOTASYPFN--MFATWSPKARLHL 1351
DB 252 RCGTTLFELLGELNCTEPLGLKNDTIPNKQITASSYTKTWGLSAFWSWPFYARLDN 311
OY 1352 QGRSNARPOVNNKEMLOVDFOKTMVGTGQVKSILTSMYKFEPLISSODGQWLF 1411

```



DB 312 OCKFNAMTAQTSASAWLQIDLSQKRVYGIITOGARDGHIOYVAAYRVAGDDGYTWT 371

OY 1412 LEFONG--KVKYFOGNDSEFTPVNSLDPPLRLRYLRHPQSWVHQIALRMEVLCG 1465

DB 372 EKXDPGASSKIFPGMNDNNSHKKNLEFPEFQARFVRIGPVAMHNRTLLVELLGC 427

RESULT 13

MFQM\_HUMAN STANDARD; PRT; 387 AA.

AC 008431;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update)

DE LACTADHERIN PRECURSOR (MILK FAT GLOBULE-EGF FACTOR 8) (MFG-EB) (HMF8)

DE (BREAST EPITHELIAL ANTIGEN BA46) (MFGM) [CONTAINS: MEDIN].

GN MFGEB.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE-Breast, and Breast carcinoma;

RX MEDLINE=96213908; PubMed=8639264;

RA Couto J.R., Taylor M.R., Godwin S.G., Ceriani R.L., Peterson J.A.;

RT "Cloning and sequence analysis of human breast epithelial antigen BA46 reveals an RGD cell adhesion sequence presented on an epidermal growth factor-like domain.";

RT DNA Cell Biol. 15:281-286(1996).

RL [2]

RN SEQUENCE OF 170-387 FROM N.A.

RP TISSUE-Mammary gland;

RC MEDLINE=91371351; PubMed=1909932;

RA Larocca D., Peterson J.A., Urra R., Kuniyoshi J., Bistrain A.M.,

RT "Isolation and characterization of full and truncated forms of human breast carcinoma protein BA46 from human milk fat globule membranes.";

RT J. Protein Chem. 17:143-148(1998).

RL [4]

RN SEQUENCE OF 268-317, AND IDENTIFICATION OF MEDIN.

RP MEDLINE=99342076; PubMed=10411933;

RA Haegggvist B., Naejlund J., Sletten K., Westermarck G.T., Mucchiano G.,

RA Thernberg L.O., Nordstedt C., Engstrom U., Westermarck P.;

RT "Medin: an integral fragment of aortic smooth muscle cell-produced lactadherin forms the most common human amyloid.";

RT Proc. Natl. Acad. Sci. U.S.A. 96:8659-8674(1999).

RL [5]

RN CHARACTERIZATION.

RP MEDLINE=97405885; PubMed=92260929;

RA Taylor M.R., Couto J.R., Scallan C.D., Ceriani R.L., Peterson J.A.;

RT "Lactadherin (formerly BA46), a membrane-associated glycoprotein expressed in human milk and breast carcinomas, promotes Arg-Gly-Asp (RGD)-dependent cell adhesion.";

RT DNA Cell Biol. 16:861-869(1997).

RL [6]

CC -1- FUNCTION: MAY BE INVOLVED IN PHOSPHOLIPID BINDING. BINDS SPECIFICALLY TO ROTAVIRUS AND INHIBITS ITS REPLICATION.

CC -1- FUNCTION: MEDIN IS THE MAIN CONSTITUENT OF AORTIC MEDIAL AMYLOID.

CC -1- SUBCELLULAR LOCATION: PERIPHERAL MEMBRANE PROTEIN.

CC -1- TISSUE SPECIFICITY: MAMMARY EPITHELIAL CELL SURFACES AND AORTIC MEDIA. OVEREXPRESSED IN SEVERAL CARCINOMAS.

CC -1- PTM: MEDIN HAS A RAGGED N-TERMINUS WITH MINOR SPECIES STARTING AT AMINO ACID 264 AND 273.

CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.

CC -1- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb.ch/announce/> or send an email to [license@isb-slb.ch](mailto:license@isb-slb.ch)).

CC EMBL: U58516; AAC50549.1; -

DR EMBL: S56151; AAB19771.1; -

DR MIM: 602281; -

DR InterPro: IPR000421; -

DR InterPro: IPR000561; -

DR Pfam: PF00008; EGF\_1.

DR Pfam: PF00754; F5\_P8\_type\_C\_2.

DR PROSITE: PS00022; EGF\_1; 1.

DR PROSITE: PS01186; EGF\_2; 1.

DR PROSITE: PS01285; FA58C\_1; 2.

DR PROSITE: PS01286; FA58C\_2; 2.

DR Signal: Glycoprotein; Milk; Repeat; EGF-like domain; Amyloid.

FT SIGNAL 1 23

FT CHAIN 24 387

FT CHAIN 202 387

FT CHAIN 268 317

FT DOMAIN 24 67

FT DOMAIN 70 225

FT DOMAIN 230 387

FT SITE 46 48

FT SITE 27 38

FT DISULFID 27 38

FT DISULFID 32 55

FT DISULFID 57 66

FT DISULFID 70 225

FT DISULFID 212 216

FT DISULFID 230 387

FT CARBOHYD 238 238

FT CARBOHYD 325 325

FT CARBOHYD 329 329

FT CARBOHYD 350 350

SO SEQUENCE 387 AA; 43123 MW; 22E6571DEC83782D CRC64;

Query Match 7.5%; Score 588; DB 1; Length 387;

Best Local Similarity 37.3%; Pred. No. 2.6e-31;

Matches 132; Conservative 69; Mismatches 125; Indels 28; Gaps 9;

OY 1126 EMPSKAGIMRVECLGHEHLAGMSLFLVYSKKOTPRGMSGHRDQITASG----- 1180

DB 48 DVFP-----YTCYCLKG--YAGNHC-----ETKVEPLGMEGNTANSQIAASSVRYTF 95

OY 1181 -QYGMAPKLRLHRYSGINAW--STKEPFSWIKVDLAPMIIHGKTGARKFSSLYI 1237

DB 96 LGIQHWPELARLNKRGAMVNAWTPSSNDNDPMLQVALLRRMVTGVYOGASRLASHETL 155

OY 1238 SGPITMYSLDGKKWQYTRGNSGTGLVVFEGVNDSSGKIKHNFPRLIARYLRPLTHYSI 1297

DB 156 KAKFVAYSLNGHEDF-FIDVYNKKHKEFVGNNKNAVHNLETPPEADQVLYLPTSCHT 214

OY 1298 RSLRLMELMGCDLNSCMPLGMSKAIQDQITASSYF-----TNMFATSPSKALRLDQG 1353

DB 215 ACLRLRELLGCELNGCANPLGKLNNSIPDKQITASSYKTWGLJHLE-SNMPYARLDDQOG 273

OY 1354 RSNAPVYNNPKEWLOVDFOKTKMYGTGVTOGVSRLSLSMVKPEFLISSQDGHQWTLF 1413

DB 274 NFNAMWAGSTGNDQWLOVDGSSKEVYGLITOGARFVSVOFVASTKVAYSNDSNMREY 333

OY 1414 F--QNGKVVYFOGNDSEFTPVNSLDPPLRLRYLRHPQSWVHQIALRMEVLCG 1465

DB 334 QDRGTSSKIFPGMNDNNSHKKNLEFPEFQARFVRIGPVAMHNRTLLVELLGC 387

RESULT 14



NR2\_HUMAN  
ID NR2\_HUMAN STANDARD: PRT: 931 AA.  
AC 060462; 014820; 014821;  
DT 01-OCT-2000 (Rel. 40, Created)  
DT 01-OCT-2000 (Rel. 40, Last sequence update)  
DT 01-OCT-2000 (Rel. 40, Last annotation update)  
DE NEUROPILIN-2 PRECURSOR (VASCULAR ENDOTHELIAL CELL GROWTH FACTOR 165 RECEPTOR 2).  
GN NR2 OR VEGF165R2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORMS A0 AND A17).  
RX MEDLINE=97470888; PubMed=9331348.  
RA Chen H., Chedotal A., He Z.-G., Goodman C.S., Tessier-Lavigne M.;  
RT "Neuropilin-2, a novel member of the neuropilin family, is a high  
RT affinity receptor for the semaphorins Sema E and Sema IV but not Sema  
RT III.";  
RL Neuron 19:547-559(1997).  
RN [2]  
RP SEQUENCE FROM N.A. (ISOFORM A22).  
RC TISSUE=Breast;  
RX MEDLINE=98188099; PubMed=9529250;  
RA Soker S., Takashima S., Miao H.-O., Neufeld G., Klagsbrun M.;  
RT "Neuropilin-1 is expressed by endothelial and tumor cells as an  
RT isoenzyme-specific receptor for vascular endothelial growth factor.";  
RL Cell 92:735-745(1998).  
RN [3]  
RP CHARACTERIZATION.  
RX MEDLINE=20309748; PubMed=10748121;  
RA Gluzman-Poltorak Z., Cohen T., Herzog Y., Neufeld G.;  
RT "Neuropilin-2 and neuropilin-1 are receptors for the 165-amino acid  
RT form of vascular endothelial growth factor (VEGF) and of placenta  
RT the 145-amino acid form of VEGF.";  
RL J. Biol. Chem. 275:18040-18045(2000).  
CC -1- FUNCTION: HIGH AFFINITY RECEPTOR FOR SEMAPHORINS 3C, 3F, VEGF-165  
CC AND VEGF-145 ISOFORMS OF VEGF. AND THE PLEK-2 ISOFORM OF PEG.  
CC -1- SUBUNIT: NEUROPILIN-2 PROBABLY FORMS AN HETEROMERIC COMPLEX WITH  
CC NEUROPILIN-1 IN ORDER TO BE A FUNCTIONAL SEMAPHORIN 3C RECEPTOR.  
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS; A22 (SHOWN HERE), A0 AND A17;  
CC ARE PRODUCED BY ALTERNATIVE SPLICING.  
CC -1- SIMILARITY: BELONGS TO THE NEUROPILIN FAMILY.  
CC -1- SIMILARITY: CONTAINS 2 CUB DOMAINS.  
CC -1- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.  
CC -1- SIMILARITY: CONTAINS 1 MAM DOMAIN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
DR EMBL; AF022859; AAC51788.1;  
DR EMBL; AF022860; AAC51789.1;  
DR EMBL; AF016098; AAC12922.1;  
DR MIM; 602070;  
DR InterPro; IPR000421;  
DR InterPro; IPR000859;  
DR InterPro; IPR000998;  
DR Pfam; PF00431; CUB; 2.  
DR Pfam; PF00629; MAM; 1.  
DR Pfam; PF00754; F5\_F8\_type\_C; 2.  
DR PROSITE; PS01180; CUB; 2.  
DR PROSITE; PS01285; FAS8C\_1; 2.  
DR PROSITE; PS01286; FAS8C\_2; 2.  
DR PROSITE; PS50060; MAM\_2; 1.  
KW Transmembrane; Glycoprotein; Neurone; Signal; Repeat; Receptor;

Alternative splicing.  
KW SIGNAL 1 20 OR 22 (POTENTIAL).  
FT CHAIN 21 931 NEUROPILIN-2. (POTENTIAL).  
FT DOMAIN 21 864 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 865 889 POTENTIAL.  
FT DOMAIN 890 931 POTENTIAL.  
FT DOMAIN 890 931 CYTOPLASMIC (POTENTIAL).  
FT DOMAIN 149 142 CUB 1.  
FT DOMAIN 149 267 CUB 2.  
FT DOMAIN 277 427 F5/8 TYPE C 1.  
FT DOMAIN 434 592 F5/8 TYPE C 2.  
FT DOMAIN 642 802 MAM.  
FT DOMAIN 671 674 POLY-SER.  
FT DISULFID 28 55 BY SIMILARITY.  
FT DISULFID 83 105 BY SIMILARITY.  
FT DISULFID 149 175 BY SIMILARITY.  
FT DISULFID 208 230 BY SIMILARITY.  
FT DISULFID 277 427 BY SIMILARITY.  
FT DISULFID 434 592 N-LINKED (GLCNAc. . .) (POTENTIAL).  
FT CARBOHYD 152 152 N-LINKED (GLCNAc. . .) (POTENTIAL).  
FT CARBOHYD 157 157 N-LINKED (GLCNAc. . .) (POTENTIAL).  
FT CARBOHYD 629 629 N-LINKED (GLCNAc. . .) (POTENTIAL).  
FT CARBOHYD 839 839 N-LINKED (GLCNAc. . .) (POTENTIAL).  
FT VARSPLIC 809 813 MISSING (IN ISOFORM A17).  
FT VARSPLIC 809 830 MISSING (IN ISOFORM A0).  
FT CONFLICT 602 602 E -> K (IN REF. 1).  
SQ SEQUENCE 931 AA; 104830 MW; 270CBABE69A0A797C CRC64;

Query Match 6.0%; Score 469.5; DB 1; Length 931;  
Best Local Similarity 28.5%; Pred. No. 5.7e-23;  
Matches 166; Conservative 87; Mismatches 185; Indels 145; Gaps 32;

QY 967 DEFDCR-AMATFSVDVLEK---DVHSLGLPLVCHNTL-----NPAHQVTVQVEFA 1016  
DB 79 EKHDKYFIEIRDBDSADLLGRKHCNIAPTTISSGMYIKFTSDYARQA--GFS 136  
QY 1017 LEFTFEDKSWYFTENNERNCRAPCNIQMEDPTREKNYRFAANGYIMDTLPLGLVMAQD 1076  
DB 137 LRYELF-KTGS-----EDSKNTSP-NGTIESGPPEKYP----- 170  
QY 1077 QRIRWYLLSMGSENHISHSFGHVFYAKKEEYKMALYNLPGVEFVEMLPKAG--- 1133  
DB 171 -----HNIDCFETILAKPKMEIIOFL--LFD-LEHDPLOYEGD 207  
QY 1134 -----IW-----RVCLIGEH-----LHAGNST-----LFIV 1155  
DB 208 KYDMDLMDWDGIPHVGPPLIGKYGCTPSELSSTGILSLFHTMVAVAKDGFSAARYYL 267  
QY 1156 YSN-----KCTPLGMASGHIRDFOITASGOY--GOMAPKILRLHYSGSINAW-----ST 1203  
DB 268 HOEPLNEOCNVPLOMGESGRINAEQISASTYSDDGRWTPQSRRL--GDDNGWTRNLDSN 325  
QY 1204 KEFPMIKVLDLAPMIHGKIQTGA--RQKFSLYISOFITIMYSLDGKWKQYTRGNSTGT 1261  
DB 326 KE---YLOVDRFLMTLMAIATOGAISRETQGYVVKYLEVSTNGEDMMVYRRGKNH-- 381  
QY 1262 LMVFGVNDSSGIRKININPPIARIYIRLHPHYISIRSTRRELMLGCDLNS--CGMPLGM 1319  
DB 382 -KVFQANNDAEEVVLNKLHAPLTLRFVRIRPQTMHSGIALRLDELGCRTADPACSNMIG 440  
QY 1320 ESKAISDAQITASSYFTNMVATWSPSKARLHLOGRSNMA--RPQVNNPKRELQVDFQKT 1376  
DB 441 LSGLIADSOISASS--TOEY-LMSSSAARL-VSSNS-GWFPRIPOAQGEEMLOYDVLGTP 495  
QY 1377 MKVGVGTQG-----VKSLLTSMYKBEFLISSODGQHWTLFPONGKV---KVFQNGND 1427  
DB 496 KTVKGVIIQARGGDSITAVEARVAFKFKVYSYLSNGDWE-YIDDPRTQPKPLEGMNH 554  
QY 1428 SETPVVNSLDPPLLTRYLRIPHQSVHO--IALRMVYVLCCEAD 1469  
DB 555 YDTPDIRRED--PIPAQVYVYPERMWSPIGMRLEVLGCDWTD 596



RESULT 15  
 NRP2\_RAT ID NRP2\_RAT STANDARD: PRT: 925 AA.  
 AC 035276;  
 DT 01-OCT-2000 (Rel. 40, Created)  
 DT 01-OCT-2000 (Rel. 40, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE NEUROPEPTIN-2 PRECURSOR (VASCULAR ENDOTHELIAL CELL GROWTH FACTOR 165 RECEPTOR 2).  
 GN NRP2.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=SPRAGUE-DAWLEY;  
 RX MEDLINE=97433085; PubMed=9288754;  
 RA Kojodkin A.L., Levengood D.V., Rowe E.G., Tal Y.-T., Giger R.J., Glnty D.D.;  
 RA "Neuropilin is a semaphorin III receptor.";  
 RL Cell 90:753-762(1997).  
 CC -1- FUNCTION: HIGH AFFINITY RECEPTOR FOR SEMAPHORINS 3C, 3F, VEGF-165 AND VEGF-145 ISOFORMS OF VEGF, AND THE PLGF-2 ISOFORM OF PGF.  
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
 CC -1- TISSUE SPECIFICITY: FOUND IN CERTAIN NEURONAL POPULATIONS OF THE CNS AND IN OTHER NONNEURONAL TISSUES INCLUDING MESENCHYMAL TISSUE LINING IN THE RIBS.  
 CC -1- SIMILARITY: BELONGS TO THE NEUROPELIN FAMILY.  
 CC -1- SIMILARITY: CONTAINS 2 CUB DOMAINS.  
 CC -1- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.  
 CC -1- SIMILARITY: CONTAINS 1 MAM DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AF016297; AAC53338.1; -.  
 DR InterPro: IPR000421; -.  
 DR InterPro: IPR000859; -.  
 DR InterPro: IPR000998; -.  
 DR Pfam: PF00431; CUB; 2.  
 DR Pfam: PF00629; MAM; 1.  
 DR Pfam: PF00754; F5\_F8\_type\_C; 2.  
 DR PROSITE: PS01180; CUB; 2.  
 DR PROSITE: PS01285; FA58C\_1; 2.  
 DR PROSITE: PS01286; FA58C\_2; 2.  
 DR PROSITE: PS50060; MAM\_2; 1.  
 KW Transmembrane; Glycoprotein; Neurone; Signal; Repeat; Receptor.  
 FT SIGNAL 1 22  
 FT CHAIN 23 925  
 FT DOMAIN 23 858  
 FT TRANSMEM 859 883  
 FT DOMAIN 884 925  
 FT DOMAIN 28 142  
 FT DOMAIN 149 267  
 FT DOMAIN 277 427  
 FT DOMAIN 434 592  
 FT DOMAIN 642 802  
 FT DISULFID 83 105  
 FT DISULFID 28 55  
 FT DISULFID 149 175  
 FT DISULFID 208 230  
 FT DISULFID 277 427  
 FT DISULFID 434 592  
 FT CARBOHYD 152 157  
 FT CARBOHYD 157 157  
 FT CARBOHYD 629 629  
 FT CARBOHYD 833 833

FT CARBOHYD 834 834 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 925 AA; 103896 MM; 3BF62903F644851C CRC64;  
 Query Match 5.9%; Score 464.5; DB 1; Length 925;  
 Best Local Similarity 28.6%; Pred. No. 1,2e-22;  
 Matches 167; Conservative 85; Mismatches 186; Indels 145; Gaps 33;  
 QY 967 DEFDCK-AMAYFSDVDLEK----DVHSGILGPLVCHTNTL-----NPAHGQVYQVEFA 1016  
 DB 79 EKHDCKYDFEIRDGDSSESADLGRKCGNAPRTIISGSLYKIFTSYAKQAQ--GFS 136  
 QY 1017 LEFTTFDETKSWYFTENMERNCRAPCNIQMEDPTEKERYRFAINQYIMDTLPLGLVMAQD 1076  
 DB 137 LREYIF-KTGS---EDCSKNFTSP--NGTIESGPFEKYP----- 170  
 QY 1077 QRIRWLLSGNSNENHSHFSGHVTYVKKEXEYKALNLYPGVEYEMLPKAG--- 1133  
 DB 171 -----HNLDCTFTLAKPRMEIILQFL--TFD-LEHDPLOYEGSD 207  
 QY 1134 -----IW-----RVECLIGE-----LHAGMST-----LELY 1155  
 DB 208 CKYDMDIDWDGIRHVGRLIKYGTGTPSKLRSSGILSLTFHTDMAVAKDGFSAHYLY 267  
 QY 1156 YSN-----KQTPLGMAAGHIIRDQITASGY--GQMAPKLRLAYSGSINAM-----ST 1203  
 DB 268 HQPPENFQCNAPDLGMEGRIANEQISASTFSDGRWTFQOSRLH--GDDNGMTFNVDN 325  
 QY 1204 KEFPSTKVDLLAPMIIHIGKTGA--RQKFSLEYISQFTIWSJLDGKKWQYRGNSTGT 1261  
 DB 326 KE--YLDVRLPLMTLATATGALSRETQGYVSKYLEVSTGEDMVMYRRGKNH- 381  
 QY 1262 LMVFGNVSSGIRKNIIFNPILARYIRLPHYISIRSTLRMLMGCDLNS--CSMPLGM 1319  
 DB 382 -KVFQANNDATELYLKNKHTPLTRIRIRPQWHGIALRLIEGCRVTDAPCSNMLGM 440  
 QY 1320 ESKAISDAQITASSYFTNMFATWSPSKARLHLOGNSNAN---RPOVNNPKENLYQVFOKT 1376  
 DB 441 LSGLIADTQISASS--TREY-LMSPSARL-VSSRS-GWFPNPPQAQPEEWLYQVLDGTP 495  
 QY 1377 MKVTGYVTQGVK--SLITSK----YKFEFLISSODGHQWTLFPQNGKV---KVFQGNOD 1427  
 DB 496 KTVKGVYIIGARGDSITLMEARAFYRKFRKVSISLGKQWE-YIDPPKQOPKLFEGNMH 554  
 QY 1428 SFTPVVNSLDPPLRLTRYLRIRHPSWVHQ--IALRMEVLCGEAOD 1469  
 DB 555 YDTPDIRREE-PVPAQYVAVYPERMSPPAGIGRLLEVLCGDWMD 596

Search completed: November 17, 2001, 13:24:36  
 Job time: 140 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2001, 13:22:11 ; Search time 26.31 Seconds  
(without alignments)  
4258.942 Million cell updates/sec

Title: US-09-689-430-2

Perfect score: 7869

Sequence: 1 MJ01ETCFEFLCLRFCSA.....WVHQLRMVELGCEAQLDY 1471

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7412	94.2	2351	1 E2HU	coagulation factor
2	6347	80.7	2133	2 T42763	coagulation factor
3	6301	80.1	2319	2 A47004	coagulation factor
4	2406.5	30.6	2224	1 KFH05	coagulation factor
5	2394.5	30.4	2183	2 T42764	coagulation factor
6	2379	30.2	2211	1 KFB05	coagulation factor
7	1729	22.0	1069	1 KUHU	ferroxidase (EC 1.1.1.1)
8	1681	21.4	1059	1 A35210	ferroxidase (EC 1.1.1.1)
9	1104	14.0	216	2 A44258	factor VIII-associ
10	663	8.4	427	2 JC4915	agg protein precu
11	657	8.3	463	1 A36479	milk fat globule m
12	650	8.3	409	2 T11743	glycoprotein anti
13	635	8.1	401	2 S65138	glycoprotein anti
14	635	8.1	427	2 S74211	glycoprotein anti
15	543	6.9	869	2 A25945	coagulation factor
16	443	5.6	927	2 J00948	A5 antigen precu
17	424.5	5.4	218	2 A47285	milk fat globule p
18	306.5	3.9	3133	2 S52093	hemocytin - silkwo
19	262	3.3	845	2 JC5256	antipocyte transcri
20	221	2.8	719	2 S51739	transcription repr
21	211	2.7	1283	2 T13799	neurexin IV - frui
22	207.5	2.6	737	2 T13139	hypothetical prote
23	206.5	2.6	737	2 T15615	hypothetical prote
24	193	2.5	913	2 A48280	receptor tyrosine
25	192	2.4	876	2 A49508	protein-tyrosine k
26	185.5	2.4	855	2 S42621	protein-tyrosine k
27	185.5	2.4	910	2 A53137	tyrosine kinase re
28	173.5	2.2	819	2 T48859	tyro 10 receptor k
29	171	2.2	1381	2 T31083	paranodin - rat

30	171	2.2	1385	2 T14158	neurexin IV - mous
31	150	1.9	1883	2 G82875	hypothetical prote
32	149.5	1.9	578	2 S63533	L-ascorbate oxidas
33	144.5	1.8	2925	2 T00133	RNA-directed RNA p
34	142	1.8	622	2 S62580	probable multicop
35	141	1.8	1217	2 T18209	mng protein - Bac
36	141	1.8	2401	2 T28676	rhoptly protein -
37	138	1.8	1193	2 JC4901	nontoxic-nonhemag
38	137.5	1.7	1178	2 S54073	probable membrane
39	136.5	1.7	1196	2 S64630	botulinum neurotox
40	135.5	1.7	1196	2 J01467	toxin, nontoxic co
41	135	1.7	343	2 T35030	probable copper ox
42	134.5	1.7	567	2 T44928	L-ascorbate oxidas
43	134.5	1.7	791	2 T16031	hypothetical prote
44	132.5	1.7	1132	2 H82887	hypothetical prote
45	131.5	1.7	580	2 F84828	probable laccase (

## ALIGNMENTS

RESULT 1  
E2HU  
coagulation factor VIII precursor [validated] - human  
N:Alternate names: antihemophilic factor A; coagulation factor VIIIc; procoagulant co  
C:Species: Homo sapiens (man)  
C>Date: 28-Aug-1985 #sequence\_revision 28-Aug-1985 #text\_change 08-Dec-2000  
C:Accession: I54318; A00525; I58059; A23584; A26174; A42348; A43986; S63527; S66445;  
R:Gitschler, J.; Wood, W.I.  
Hum. Mol. Genet. 1, 199-200, 1992  
A:Title: Sequence of the exon-containing regions of the human factor VIII gene.  
A:Reference number: I54318; MUID:93265012  
A:Accession: I54318  
A:Status: Preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1921, 'S', 1923-2351 <RES>  
A:Cross-references: GB:M8648; NID:g182381; PIDN:AAA52420.1; PID:g182383  
R:Wood, W.I.; Capon, D.J.; Simonsen, C.C.; Eaton, D.L.; Gitschler, J.; Keyt, B.; Seeb  
Nature 312, 330-337, 1984  
A:Title: Expression of active human factor VIII from recombinant DNA clones.  
A:Reference number: A00525; MUID:85061548  
A:Accession: A00525  
A:Molecule type: mRNA  
A:Residues: 1-2351 <MOO>  
A:Cross-references: EMBL:X01165; EMBL:X01166; EMBL:X01179  
R:Toole, J.J.; Knopf, J.L.; Wozney, J.M.; Sultzman, L.A.; Buecker, J.L.; Pittman, D.D  
s, D.N.; Hewick, R.M.  
Nature 312, 342-347, 1984  
A:Title: Molecular cloning of a cDNA encoding human antihemophilic factor.  
A:Reference number: I58059; MUID:85061550  
A:Accession: I58059  
A:Status: nucleic acid sequence not shown; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-74, 'V', 76-1259, 'E', 1261-2351 <RE2>  
A:Cross-references: GB:K01740; NID:g182802; PIDN:AAA52484.1; PID:g182803  
R:Truett, M.A.; Blacher, R.; Burke, R.L.; Caput, D.; Chu, C.; Dina, D.; Hattog, K.; K  
B.; Randojph, A.; Urdeda, M.S.; Valenzuela, P.; Dahl, H.H.; Favalato, J.; Hansen, J.;  
DNA 4, 333-349, 1985  
A:Title: Characterization of the polypeptide composition of human factor VIII:C and t  
A:Reference number: A23584; MUID:86081164  
A:Accession: A23584  
A:Molecule type: mRNA  
A:Residues: 1-2351 <TRD>  
A:Cross-references: GB:M4113; NID:g182817; PIDN:AAA52485.1; PID:g182818  
R:Eaton, D.; Rodriguez, H.; Vohar, G.A.  
Biochemistry 25, 505-512, 1986  
A:Title: Proteolytic processing of human factor VIII. Correlation of specific cleavag  
ity.  
A:Reference number: A26174; MUID:86159740  
A:Accession: A26174  
A:Molecule type: protein  
A:Residues: 20-36; 392-399, 'X', 401-402; 1668-1678; 1709-1722, 'D', 1723-1725; 1741-1755 <EA  
R:Pittman, D.D.; Wang, J.H.; Kaufman, R.J.



Biochemistry 3315-3325, 1992

A:Title: Identification and functional importance of tyrosine sulfate residues within re

A:Reference number: A42348; MUID:92207952

A:Accession: A42348

A:Molecule type: Protein

A:Residues: 20-36;356-371;392-408;582-594;1668-1669,'X',1671;1672-1692;1693-1708;1709-17

A:Experimental source: recombinant material from Chinese hamster ovary cells

A:Note: sequence extracted from NCBI backbone and corrected to correspond with the publi

R:Pay, P.J.; Smudzin, T.M.

J. Biol. Chem. 264, 14005-14010, 1989

A:Title: Intersubunit fluorescence energy transfer in human factor VIII.

A:Reference number: A43966; MUID:89340500

A:Accession: A43966

A:Molecule type: Protein

A:Residues: 'X',517-523;1853-1860,'X',1862-1864,'X',1866 <FAY>

R:LeYTE, A.; van Schlingel, H.B.; Niehrs, C.; Huttner, M.P.; Mertens, K.;

J. Biol. Chem. 266, 740-746, 1991

A:Title: Sulfation of Tyr(1680) of human blood coagulation factor VIII is essential for

A:Reference number: A56109; MUID:91093266

A:Contents: annotation: sulfation

R:Gitschler, J.; Wood, W.I.; Goralka, T.M.; Wilson, K.L.; Chen, E.Y.; Eaton, D.H.; Vehar,

Nature 312, 326-330, 1984

A:Title: Characterization of the human factor VIII gene.

A:Reference number: A56196; MUID:85061547

A:Contents: annotation: Introns

R:McMullen, B.A.; Fujikawa, K.; Davie, E.W.; Hedner, U.; Ezban, M.

Proc Natl Acad Sci U S A. 74, 740-746, 1995

A:Title: Locations of disulfide bonds and free cysteines in the heavy and light chains c

A:Reference number: A56216; MUID:9538127

A:Contents: annotation: disulfide bonds

A:Note: 329-Cys, 711-Cys, and 2019-Cys were shown to have free sulphhydryls

R:Kjølbye, M.; Hedning, A.; Talbo, G.; Persson, E.; Thomsen, J.; Ezban, M.

Eur. J. Biochem. 234, 773-779, 1995

A:Title: Amino acid residues 721-729 are required for full factor VIII activity.

A:Reference number: S63527; MUID:96163459

A:Accession: S63527

A:Molecule type: Protein

A:Residues: 733-752;753-759 <KJA>

R:Lind, P.; Larsson, K.; Spita, J.; Sydow-Baeckman, M.; Almstedt, A.; Gray, E.; Sandberg

Eur. J. Biochem. 232, 19-27, 1995

A:Title: Novel forms of B-domain-deleted recombinant factor VIII molecules. Construction

A:Reference number: S66445; MUID:96048024

A:Accession: S66445

A>Status: preliminary

A:Molecule type: Protein

A:Residues: 1668-1685 <LIN>

C:Comment: Factor VIII is activated by factor Xa and thrombin, but prolonged exposure pr

C:Genetics:

A:Gene: GDB:F8C

A:Cross-references: GDB:119124; OMIM:306700

A:Map position: Xq28-Xq28

A:Introns: 48/2; 89/1; 130/1; 201/1; 224/1; 263/1; 337/1; 424/2; 481/3; 513/1; 584/3; 65

C:Function:

A:Description: acts as a cofactor, with calcium and phospholipid, for the factor IXa prot

A:Pathway: blood coagulation

C:Superfamily: coagulation factor VIII; discoidin I amino-terminal homology; ferroxidase

C:Keywords: acute phase; blood coagulation; duplication; glycoprotein; hemophilia A; pla

F:1-19/Domin: signal sequence #status predicted <SIG>

F:20-2351/Product: coagulation factor VIII #status experimental <MAT>

F:20-740/Product: coagulation factor VIIIA heavy chain #status experimental <ACH>

F:20-356/Domin: A1 <A1>

F:23-348/Domin: ferroxidase repeat homology <FOI>

F:392-759/Domin: A2 <DA2>

F:402-730/Domin: ferroxidase repeat homology <FO2>

F:760-1667/Domin: B <DBO>

F:1668-2351/Product: coagulation factor VIIIA light chain #status experimental <ACL>

F:1709-2038/Domin: A3 <DA3>

F:1716-2038/Domin: ferroxidase repeat homology <FO3>

F:2039-2191/Domin: C1 <DC1>

F:2039-2188/Domin: discoidin I amino-terminal homology <DN1>

F:2192-2351/Domin: C2 <DC2>

F:2192-2345/Domin: I amino-terminal homology <DN2>

F:60,258,601,776,803,847,919,962,982,1020,1024,1074,1085,1204,1274,1278,1301,1319,1403,1

[illegible]



D	b	841	PGALDSNNLSLEMHFRPQLHHSDMVFTPEESGLQLRLNEKLTATATELKKLDRKVSST	900
Q	y	779	-----	778
D	b	901	SNNLISTIPSDNLAAGTDNNTSSLGPESPMAYHDSQLDITTLFGKSSPLTESGPLSLSEE	960
Q	y	779	-----	778
D	b	961	NNDSKLLESGLMNSQESSMCKNVSSTESGRLEFGKRAHGALLTKDNALFKVISILAKTN	1020
Q	y	779	-----	778
D	b	1021	KTSNNSATNKRTHIDGSSLIEHNSPSWONILESDEFEKKVPIPLIHRLMLDKNATALRL	1080
Q	y	779	-----	778
D	b	1081	NHMSNKTTSKKNMEMVOQKKKEGPIRPDAQNPDMSPFKMLFLPESARWIOETHGKNSLNSG	1140
Q	y	779	-----	778
D	b	1141	QGPSPKOLVSLGPEKSVEGONFLSEKNKVVVGKEPTKDVGLKENVPSSRMLFLTNDN	1200
Q	y	779	-----	778
D	b	1201	LHENNTNNOEKKIOELEKEKETLQENVVLPOLIHVTYGIRKNPKMNLFLISTQONVEGSTD	1260
Q	y	779	-----	778
D	b	1261	GAYAPVLQDFRSLNDSTNRTKHTAHFSKGBEENLEGNOTKQLEKXACTRISPNT	1320
Q	y	779	-----	778
D	b	1321	SQONFVQSRKRALKQERLPLEETELEKRIYDDTSTONSKNMKHLTPSTLQIDYNEKE	1380
Q	y	779	-----	778
D	b	1381	KGATIGPSLDCLRHSIFQANRSPILAKVSSPFSIRPIYLTRVLFPDNSHLPAASY	1440
Q	y	779	-----	778
D	b	1441	RKKGVOESSHFLQAGAKNNLSLAILTLEMTGDQREVSLGTSATNSVYKKEVENTVLP	1500
Q	y	779	-----	778
D	b	1501	KPDLPKTISGVKVELLPKHVITYQKDLFPTETSNGPSGHLDIVESGLLOGTGAIKMEANRP	1560
Q	y	779	-----	778
D	b	1561	GKVPFLRVATESAKTPSKLLDPLAMDNHYGTQIPKEEMKSOEKSPKTAFAKKDITL	1620
Q	y	779	-----PPVLRKHQREITRTTLOSODE	800
D	b	1621	NACESNHAIALINEGONKPEIEVMTMAKOGTERLCSQNPVPLKRHQREITRTTLOSODE	1680
Q	y	801	IDYDDTISVEKKEDPIYEDENOSPRFOKTRHAYFLAAVERLMDYCMSSPVLNRR	860
D	b	1681	IDYDDTISVEKKEDPIYEDENOSPRFOKTRHAYFLAAVERLMDYCMSSPVLNRR	1740
Q	y	861	AQSGSVQPKFVVFOEFTDGSFTOPLYRGELENHGLGPIYRAEYEDNIWTFPRNQASR	920
D	b	1741	AQSGSVQPKFVVFOEFTDGSFTOPLYRGELENHGLGPIYRAEYEDNIWTFPRNQASR	1800
Q	y	921	PYSFYSLSLIEYEDQOGAPBRKNFYKPNETKTFPKKVOHNAAPTRKDEPDCAMAYFSDV	980
D	b	1801	PYSFYSLSLIEYEDQOGAPBRKNFYKPNETKTFPKKVOHNAAPTRKDEPDCAMAYFSDV	1860
Q	y	981	DLEADVHSGLIGPLLYCHTNTLMPAHRGROYVQEFALFTTIDETKSWTFETNMERNORA	1040
D	b	1861	DLEADVHSGLIGPLLYCHTNTLMPAHRGROYVQEFALFTTIDETKSWTFETNMERNORA	1920
Q	y	1041	PCNTOMEDPTFKENYRPHALINGTIYMDTLPGVLVAQDQIRWYLLSGNSNENIHSHFSGH	11000

Db 1921 PCNIQMEDPTFEKNRFRPAHNGYIMDTLPGLVMAODQIRMYLLSMGSENMHSIHFSGH 1980

Qy 1101 VFTVRKKEEYKKALNLKPVGEIVEMLPSKAGIMRWCLIGELHAGMSTPLFVYSNKC 1160  
|||||  
Db 1981 VFTVRKKEEYKKALNLKPVGEIVEMLPSKAGIMRWCLIGELHAGMSTPLFVYSNKC 2040

Qy 1161 QTPPLMASGHIRFOITASGOTGWAPKLARLHSGSINAMSTEPEFSWIKVDLAPMI 1220  
|||||  
Db 2041 QTPPLMASGHIRFOITASGOTGWAPKLARLHSGSINAMSTEPEFSWIKVDLAPMI 2100

Qy 1221 HKIKTGAKRKRFSSLYISQFIIMYSLDGKKMOYRGNSGTGLMPFGGVNDSSGIKHNFEN 1280  
|||  
Db 2101 HKIKTGAKRKRFSSLYISQFIIMYSLDGKKMOYRGNSGTGLMPFGGVNDSSGIKHNFEN 2160

Qy 1281 PPIIARYRLRHPTHYSIRSTLRMEIMGCDLNSCMSPMLMESKAISDAOITASSEYTNMFA 1340  
|||||  
Db 2161 PPIIARYRLRHPTHYSIRSTLRMEIMGCDLNSCMSPMLMESKAISDAOITASSEYTNMFA 2220

Qy 1341 TWSPSKARLHLQGRSNAMPPOVNNPKEWLOVDFOKTMKVTVGTTQGVKSLLTMYVEFL 1400  
|||||  
Db 2221 TWSPSKARLHLQGRSNAMPPOVNNPKEWLOVDFOKTMKVTVGTTQGVKSLLTMYVEFL 2280

Qy 1401 ISSSDQGHQMTLFPPONGKVKVFCQGMDSFPTVVNSLDPELLTRLRLRHPSGWVHQIALRM 1460  
|||||  
Db 2281 ISSSDQGHQMTLFPPONGKVKVFCQGMDSFPTVVNSLDPELLTRLRLRHPSGWVHQIALRM 2340

Qy 1461 EYLCEADQDLY 1471  
|||||  
Db 2341 EYLCEADQDLY 2351

RESULT 2  
T42763  
coagulation factor VIII precursor - pig  
C.Species: Sus scrofa domestica (domestic pig)  
C.Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 09-Jun-2000  
C.Accession: T42763  
R.Lollar, P.  
submitted to the EMBL Data Library, August 1996  
A.Reference number: Z22269  
A.Accession: T42763  
A.Status: preliminary; translated from GB/EMBL/DDBJ  
A.Molecule type: mRNA  
A.Residues: 1-2133 <OL>  
A.Cross-references: EMBL:U49517; NID:g15l1633; PID:g15l1634; PIDN:AAB06705.1  
C.Superfamily: coagulation factor VIII; discoidin I amino-terminal homology; ferroxid C.Keywords: acute phase; blood coagulation; duplication; glycoprotein; hemophilia A; F.1-19/Domain: signal sequence #status predicted <SIG>  
F.20-213/Product: coagulation factor VIII #status predicted <MAT>  
F.23-349/Domain: ferroxidase repeat homology <FOX1>  
F.403-730/Domain: ferroxidase repeat homology <FOX2>  
F.1498-1820/Domain: ferroxidase repeat homology <FOX3>

Query Match 80.7%; Score 6347; DB 2; Length 2133;  
Best Local Similarity 58.0%; Pred. No. 0;  
Matches 1237; Conservative 107; Mismatches 126; Indels 664; Gaps 4;

Qy 1 MOLEISTFFPICLRFCCSATRRYYLGAVELSDWTMOD-LIGELPVDARRPPRPVSKSPF 59  
|||||  
Db 1 MOLEISTVFCLCPPLGFSAIRRYLGAVELSWDPYRSELLRELVHTREPATPGALPL 60

Qy 60 NTMSVYKKTTLVEEPVLHFNIAKP RPMMGLGPTIOAEVYDVVTILKNMASHPVSLHA 119  
|||||  
Db 61 GPSVYTKTTFVEFPDLQLESVARPPPMWGGLGPTIOAEVYDVVTILKNMASHPVSLHA 120

Qy 120 VGVSYWKASEAEVDDQTQSREKEDDKVFPGSHTYVOVLKENGPMASDPLCTYSLS 179  
|||||  
Db 121 VGVSYWKASEAEVDDHTSQREKEDDKVLPKSOITYWOVLKENGPMASDPLCTYSLS 180

Qy 180 HVDLVKDNSGLIGALLVCREGSLAKEKTOTLIHKFILLFAVEDGSKMHSSETKNSLMODR 239  
|||||  
Db 181 HVDLVKDNSGLIGALLVCREGSLAKEKTOTLIHKFILLFAVEDGSKMHSSETKNSLMODR 240



[illegible]

Qy	779	-----	778
Db	1320	KVRVHREDLEÖKTSNVSACAHGDLGOEFLÖKTRGPVNLKVRNRPRTSPKLLGPPMPKE	1375
Qy	779	-----	778
Db	1380	WESLEKSPKSTALTKDILSLPDRHSSNHSIAKKNCGAETPQREAAWTKÖGGPGRCLAP	1433
Qy	779	-PVLKRÖHEITRTTLÖSDÖEEIDYDITISVEMKKEFDIYDENÖSPRSFÖKTRHY	837
Db	1440	KPVLRRHÖRDISPTPOPEDKKMDYDIFSTETKGEDFIEYEDENODPRSÖKTRHY	1499
Qy	838	FLAANERLMDXGMSSPHVLBNRQÖSVQÖFKKVVQÖFTÖSFLQPLRYCELNPHGL	897
Db	1500	FLAAEÖLMDXGMSPPALNRQÖNEVPFRKKVVFREPADÖSFPÖSTRGELNHJL	1555
Qy	898	LGPLYRAEVEDINWTFERNQASRPYSYSSLISYEEDORGAEPRTNPFVKPNETKTEYWK	957
Db	1560	LGPLYRAEVEDINWTFERNQASRPYSYSSLISYPPDÖEGAPRINFPQNETRTYWK	1613
Qy	958	VÖHHPAPPKDEFÖCKANARYSDVDLEKDVSGLGLPLVCHTNTLBNPARGROYTYÖEFL	1011
Db	1620	VÖHHPAPPEDEFÖCKANARYSDVDLEKDVSGLGLPLLCRANTLNAARGROYTYÖEFL	1675
Qy	1018	FETTFDETKSVYPTFENNERNCRAPCNIÖMDEPFRKENVRYHAINGYIMDTLGLVMAÖQÖ	1077
Db	1660	FETTFDETKSVYPTFENNERNCRAPCHLÖMDEPFLKENYRPHAINGYIMDTLGLVMAÖQÖ	1733
Qy	1078	RIRWYLLSMGSENIHSHIFSGHFTYRKKEEYKMALYNLPGVFETVEMLPBSKAGIMRY	1133
Db	1740	RIRWYLLSMGSENIHSHIFSGHVSFRKKEEYKMAVYNLPGVFETVEMLPBSKAGIMRY	1799
Qy	1138	ECLIGEHLHAGMSTLFLVYSNKÖCTPLGMAASGHTRPÖITASÖGYÖMAPKLARLHYSGS	1197
Db	1800	ECLIGEHLHÖQMSTPLFLVYSSECCAPRGMASGRIRPÖITASÖGYÖMAPKLARLHYSGS	1855
Qy	1198	INANSTKEPFMIVDVLAPRIIHGICITQÖARÖKFSLSYÖPFIWYSLDGKMMÖTÖYGN	1257
Db	1860	INANSTKDPHMKIVDVLAPRIIHGIMTÖQARÖKFSLSYÖPFIWYSLDGKMMÖSTÖYGN	1913
Qy	1258	STGTLMVFEFGNVDSGIGKHNIFNPBITARYIRLPHPHYSIRSTLRLMELGCOLNSCÖMPL	1317
Db	1920	STGTLMVFEFGNVDSGIGKHNIFNPPIARYIRLPHPHYSIRSTLRLMELGCOLNSCÖMPL	1975
Qy	1318	GMEKSAISDAÖITASSYPTFNNFAWMSÖSKARLHÖGRSNAAMRÖVNNPKEMLÖVDPÖKTM	1377
Db	1980	GÖMÖKSAISDSÖITASSHSLSNIFAWMSÖSARLHÖGTNNAWRPVSASAEEMLÖVDPÖKTV	2035
Qy	1378	KVTGTTÖQGVNSLTTSMYKEFLISSÖDÖHÖMTLFEFONKVVYVFEÖGNDÖSTFPVNSLD	1437
Db	2040	KVTGTTÖQGVNSLTTSMYKEFLIVSSÖDÖRRMTLFLÖDÖHÖMTLVFVGNDÖSSTFPVNMALD	2099
Qy	1438	PLLTLRYLRHPÖSWHÖIALRMVNLCEÖKÖDLY 1471	
Db	2100	PPLEFTRYLRHPÖTMAÖHIALRLREVNLCEÖKÖDLY 2133	

RESULT 3  
A47004 coagulation factor VIII precursor - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 09-Sep-1994 #sequence\_revision 09-Sep-1994 #text\_change 18-Jun-1999  
C:Accession: A47004  
R:Elder, B.; Takich, D.; Gitschler, J.  
Genomics 16, 374-379, 1993  
A:Title: Sequence of the murine factor VIII cDNA.  
A:Reference number: A47004; MUID:93300511  
A:Accession: A47004  
A:Status: Preliminary  
A:Molecule type: mRNA  
A:Residues: 1-2319 <ELD>  
A:Cross-references: GB:LO5573; NID:g192456; PIDN:AAA37385.1; PID:g192457  
;Superfamily: coagulation factor VIII; discoidin I amino-terminal homology; ferroxiid

ferrooxid



F:1-19/Domain: signal sequence #status predicted <SIG>  
F:23-349/Domain: ferroxidase repeat homology <FO1>  
F:402-730/Domain: ferroxidase repeat homology <FO2>  
F:1686-2006/Domain: ferroxidase repeat homology <FO3>  
F:2007-2156/Domain: discoidin I amino-terminal homology <DN1>  
F:2160-2313/Domain: discoidin I amino-terminal homology <DN2>

Query Match 80.1%; Score 6301; DB 2; Length 2319;  
Best Local Similarity 53.8%; Pred. No. 0;  
Matches 1251; Conservative 102; Mismatches 111; Indels 862; Gaps 10;

```
QY 1 MQEISTCFECLIFRCCSATRRYLGAVELSDMQSD-IGELPVDARFPRVRSPEF 59
    ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 MQIALPACFELSLFNCSSAIRRYLGAVELSMNIQSDLSVLITDSRFLPRMSTSEF 60
QY 60 NTSVYKKTLEVEFTVHLFNIAKPRPMWGLGPPIQAEVYDTVVITLKNASHPVSLHA 119
    ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 61 NTSIMYKKTVEYKQDLFNIAKPRPMWGLGPPIQAEVYDTVVITLKNASHPVSLHA 120
QY 120 VGVSYWKASGAEVDDTSQREKEDKVPFGSHTYVWQVLKENGPMASDPLCLTYSL 179
    ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 121 VGVSYWKASGAEVDDTSQREKEDKVPFGSHTYVWQVLKENGPMASDPLCLTYSL 180
QY 180 HYDLVKDLSGLIGALLVCRGSLAKEKQTLHKFILFAVEDEKSMHSETKNSLMDR 239
    ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 181 HYDLVKDLSGLIGALLVCRGSLAKEKQTLHKFILFAVEDEKSMHSETKNSLMDR 240
QY 240 DAASARAPKMHVNGVYVNRSLPGLIGCHRSVYWHVIGMGTPEVHSIFLEGHTFLVRN 299
    ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 241 DSASARAPKMHVNGVYVNRSLPGLIGCHRSVYWHVIGMGTPEVHSIFLEGHTFLVRN 300
QY 300 HRQASLESPITFLTAQTLMDLGOFLFCHTSHOHDMEAHYVAVDSCPPEPQLRMK-N 358
    ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 301 HRQASLESPITFLTAQTLMDLGOFLFCHTSHOHDMEAHYVAVDSCPPEPQLRMK-N 360
QY 359 NEEADYDDDLTDEMDVVRFDSDNSPSFIQIRSAKKHKKPTWVYIAEEDMDYAPLV 418
    ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 361 NEEADYDDDLTDEMDVVRFDSDNSPSFIQIRSAKKHKKPTWVYIAEEDMDYAPLV 418
QY 419 LAPDRRSKSOYLNNGPQIRGKRYKRVFMAVYDEFTKREAIQIESGILGFLLYGEVD 478
    ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 419 PTDNGSKSOYLSNGPGRIGRKYKRVFIAYTDEFTKREAIQIESGILGFLLYGEVD 478
QY 479 TLLIIFKQASRPVNIYHGHTDVRPLYSRRLPKGVKHLKOPFLPGELFKYKTVYED 538
    ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 479 TLLIIFKQASRPVNIYHGHTDVRPLYSRRLPKGVKHLKOPFLPGELFKYKTVYED 538
QY 539 GPTKSDPCLTRYSSSFVNMERDLASGLIGPLLCYKESVDQRGNOIMSDKRNVLFSVF 598
    ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 539 GPTKSDPCLTRYSSSFVNMERDLASGLIGPLLCYKESVDQRGNOIMSDKRNVLFSVF 598
QY 599 DENRSWYLTENIQRLPNPAGVQLEDPFOASNMHSINGYVFDLSQVSYCLHEVAYYI 658
    ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 599 DENRSWYLTENIQRLPNPAGVQLEDPFOASNMHSINGYVFDLSQVSYCLHEVAYYI 658
QY 659 LSIQAQDFELSVFSGYFFKHKMYEDTLTPPSGGEVFMSEMPGMLIIGCHNSDRN 718
    ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 659 LSVGAQDFELSVFSGYFFKHKMYEDTLTPPSGGEVFMSEMPGMLIIGCHNSDRN 718
QY 719 RGMATLLKVSSCDKNTGYDYDSYEDISAVLLSKNNAIEPSSONSRRPSTROKOF--- 775
    ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 719 RGMATLLKVSSCDKNTGYDYDSYEDISAVLLSKNNAIEPSSONSRRPSTROKOF--- 775
QY 776 RGMATLLKVSSCDKNTGYDYDSYEDISAVLLSKNNAIEPSSONSRRPSTROKOF--- 775
    ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 776 RGMATLLKVSSCDKNTGYDYDSYEDISAVLLSKNNAIEPSSONSRRPSTROKOF--- 775
QY 779 TIRPKNMEKIEPQEEIEMLVQSVSYSDMLLIGOSHPTPHGLFLSDQEAIEALIH 838
    ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 779 TIRPKNMEKIEPQEEIEMLVQSVSYSDMLLIGOSHPTPHGLFLSDQEAIEALIH 838
QY 776 --- 775
    ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 776 --- 775
    ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 839 DHSPAIDNSNGPSKVYQLRPESHSEKIVTFPOGQLRSNKSLETTIEVKKMKLGLQV 898
    ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 839 DHSPAIDNSNGPSKVYQLRPESHSEKIVTFPOGQLRSNKSLETTIEVKKMKLGLQV 898
QY 776 --- 775
    ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
```

```
Db 899 SSLPNSLMNTTLLSDNLKATFEKTDSSGPPDMPVHSSSKLSTTARCAKAYSLVGSHPVLN 958
    ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 776 --- 775
    ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 959 ASENSDNLIDSTLMYSQESLPDRNLTLENDRLREKRFHIALTKDNTLFRKDNVSL 1018
    ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 776 --- 775
    ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1019 MKNKTYNHSTTNEKLHESPTSIENSTTDDAILKYNSEIQEYVATLIDHTLLGKNST 1078
    ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 776 --- 775
    ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1079 YLRLLHMLNRTTSTKNKDLFHRKDEDPIDODENTIMPSKMLFLSESSNMFKKTNGNS 1138
    ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 776 --- 775
    ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1139 LNSEQHSPKQLVLMFKFYVKNQSLSEKNKVVEODGFTKNIGLKDMAFPHNSIFLT 1198
    ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 776 --- 775
    ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1199 TISNVHENGRIHNOEKNIOEELKEALIEKVVLPOVHEATGSKNFLKDLILGTRONISL 1258
    ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 776 --- 778
    ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1259 YEAVHPVLQNTISNTNTVQIIMHEHFFKRKKDETNSEGLVKNKTRKMYKPSQKNT 1318
    ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 779 --- 778
    ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1319 TQSKRALGQFLSTQWLKTINCSTQCIIKQIDHSKEMKFFTKSLSDSVYKSTQTN 1378
    ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 779 --- 778
    ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1379 SSDSHIVTSAPPIDLRSPQNKFSHVQASSYIDFKTKSRIOESNNFLKETINNP 1438
    ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 779 --- 778
    ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1439 SLAILPMMKFIQDGFSTPSKNTSNVYKRENIIFLKPILPESGKIELLPQVSIQEE 1498
    ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 779 --- 778
    ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1499 EILPETSHSGPGLNLKKEVLAQIQGPTKWNKAKRHGESIKGTSSKNTRSKLNH 1558
    ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 779 --- 778
    ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1559 AMDHYAAOIPKDMKSKSEKSPETIISIKQEDTILSLRPHGNSHIGANEKQMPQRETTW 1618
    ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 779 --- 825
    ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1619 VKQGTQRTCSQIRPVLRKHQREL--SAFQSEQEAETVDDATITIE-TIEDPDYSEDIKQ 1675
    ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 826 SPRSFQKTRHYFLAVERLMDYGMSSSPHYLRNAGSGSVPOFKVVFQEFDTGSGTOP 885
    ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1676 GPRSFQKTRHYFLAVERLMDYGMSSSPHYLRNAGSGSVPOFKVVFQEFDTGSGTOP 885
    ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 886 LYRGLNLHGLGPRYIAEVEDNIMTFERNQASRPYSFSLISYEDDOAGAPRRNF 945
    ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1735 LYRGLNLHGLGPRYIAEVEDNIMTFERNQASRPYSFSLISYEDDOAGAPRRNF 945
    ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 946 VKPNFTKTYFMKVQHMMAPTKDEDFCKAMAYFSVDLKKDVHSGILGLVLCHTNTLMPA 1005
    ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1794 VKPNFTKTYFMKVQHMMAPTKDEDFCKAMAYFSVDLKKDVHSGILGLVLCHTNTLMPA 1005
    ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1006 HGRQVTVQEFALFTIFDETYSWYFTENMERNCRAPCNIQMEDPTKENYRPHALINGYIM 1065
    ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1854 HGRQVTVQEFALFTIFDETYSWYFTENMERNCRAPCNIQMEDPTKENYRPHALINGYIM 1065
    ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1066 DTLGGLVMAODORTIRMYLLSGNSNENIHSIHSGVFTVRKKEEKYKMLVNLVPEVFTY 1125
    ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1914 DTLGGLVMAODORTIRMYLLSGNSNENIHSIHSGVFTVRKKEEKYKMLVNLVPEVFTY 1125
    ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1126 EMLPSKAGIWRVECLIGELHLHAGNSTLFLVYSNKCQPLGNASGHIRDFQTLASGOYQW 1185
    ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
```



Db 1974 EMIPBAGIWEVCELIQHEHLDQAGMSTFLVYSQCOPIPLGMASSGTRDQITASGHYQW 2033  
Qy 1186 AFKLARLHSGSINAMSTKEPFSWIKVDLAPMIIGIKTQGAROKFSSLYISOFITMYS 1245  
Db 2034 APNLARLHSGSINAMSTKEPFSWIKVDLAPMIIGIKTQGAROKFSSLYISOFITMYS 2093  
Qy 1246 LDGKRWQYIRGNSGTGLTLMVFEQVNDSSGKIKHNIFNPPIIARYRLPHPTHSISITLMEL 1305  
Db 2094 LDGKRWQYIRGNSGTGLTLMVFEQVNDSSGKIKHNIFNPPIIARYRLPHPTHSISITLMEL 2153  
Qy 1306 MGCDSNCSGPIGMEKSKATSDAQITASSTFTNMFATWSPKARLHLOGRSNARPOVNNP 1365  
Db 2154 MGCDSNCSGPIGMEKSKATSDAQITASSTFTNMFATWSPKARLHLOGRSNARPOVNNP 2213  
Qy 1366 KEMLOVDFOKTKMKVGVTTQGVKSLTSMYKVEKFLISSQDGHQMTLFFQNGKVKVQGN 1425  
Db 2214 KEMLOVDFOKTKMKVGVTTQGVKSLTSMYKVEKFLISSQDGHQMTLFFQNGKVKVQGN 2273  
Qy 1426 QDSFTPVVNSLDPRPLTFRYLHPOGSWVHOIALRMEVLGCEADOLY 1471  
Db 2274 QDSFTPVVNSLDPRPLTFRYLHPOGSWVHOIALRMEVLGCEADOLY 2319

## RESULT 4

KFH05  
coagulation factor V precursor [validated] - human  
N:Alternate names: coagulation labile factor; proaccelerin  
C:Species: Homo sapiens (man)  
C:Date: 19-May-1989 #sequence revision 02-Jun-1995 #extl\_change 08-Dec-2000  
C:Accession: A561172; A42344; A28028; A27498; A25887  
R:Cltipe, L.D.; Moore, K.D.; Kane, W.H.  
Biochemistry 31, 3777-3785, 1992  
A:Title: Structure of the gene for human coagulation factor V.  
A:Reference number: A42344; MUID:9223268  
A:Accession: A561172  
A:Molecule type: DNA  
A:Residues: 1-2224 <CRI>  
A:Cross-references: GB:J05368  
A:Accession: A42344  
A:Molecule type: DNA  
A:Residues: 48-58; 79-89; 120-130; 191-201; 239-249; 313-323; 368-378; 428-437; 461-471; 533-542;  
2070; 2111-2120; 2172-2181 <CR2>  
R:Jenny, R.J.; Piltman, D.D.; Toole, J.J.; Kriz, R.W.; Aldape, R.A.; Hewick, R.M.; Kaufm  
Proc. Natl. Acad. Sci. U.S.A. 84, 4846-4850, 1987  
A:Title: Complete cDNA and derived amino acid sequence of human factor V.  
A:Reference number: A28028; MUID:87260886  
A:Accession: A28028  
A:Molecule type: mRNA  
A:Residues: 1-857; R, 859-864, 'R', 866-924, 'E', 926-1763, 'T', 1765-2212, 'T', 2214-2224 <JEN>  
A:Cross-references: GB:M16967  
A:Note: parts of this sequence, including the amino end of the mature protein, were dete  
R:Kane, W.H.; Ichinose, A.; Hagen, F.S.; Davie, E.W.  
Biochemistry 26, 6508-6514, 1987  
A:Title: Cloning of cDNAs coding for the heavy chain region and connecting region of hum  
A:Reference number: A27498; MUID:88107560  
A:Accession: A27498  
A:Molecule type: mRNA  
A:Residues: 1-1284; 'I', 1286-1600 <KAN>  
A:Cross-references: GB:M1785  
A:Note: parts of this sequence were determined by protein sequencing  
R:Kane, W.H.; Davie, E.W.  
Proc. Natl. Acad. Sci. U.S.A. 83, 6800-6804, 1986  
A:Title: Cloning of a cDNA coding for human factor V, a blood coagulation factor homolog  
A:Reference number: A25897; MUID:86313665  
A:Accession: A25897  
A:Molecule type: mRNA  
A:Residues: 1188-1215, 1315-2224 <KA2>  
A:Cross-references: GB:M14335  
A:Note: parts of this sequence were determined by protein sequencing  
R:Keller, F.G.; Otefl, T.L.; Quinn-Allen, M.A.; Kane, W.H.  
Biochemistry 34, 4118-4124, 1995  
A:Title: Thrombin-catalyzed activation of recombinant human factor V.  
A:Reference number: A56139; MUID:95210278

A:Contents: annotation; thrombin cleavage sites  
C:Comment: Factor V is activated by thrombin and partially by coagulation factor Xa.  
C:Genetics:  
A:Gene: GDB:F5  
A:Cross-references: GDB:119896; OMIM:227400  
A:Map position: 1q23-1q23  
A:Intons: 53/2; 84/1; 125/1; 196/1; 244/1; 318/1; 373/2; 432/3; 466/1; 537/3; 588/1;  
C:Function:  
A:Description: acts as a cofactor, with calcium and phospholipid, for the factor Xa p  
A:Pathway: blood coagulation  
C:Superfamily: coagulation factor V; discoidin I amino-terminal homology; ferroxidase  
C:Keywords: blood coagulation; duplication; glycoprotein; phospholipid binding; plasm  
F:1-28/Domain: signal sequence #status predicted <Sig>  
F:29-2224/Product: coagulation factor V #status predicted <Mat>  
F:29-737/Product: coagulation factor Va heavy chain #status experimental <VAV>  
F:29-345/Domain: A1 <DA1>  
F:33-329/Domain: ferroxidase repeat homology <FO1>  
F:346-691/Domain: A2 <DA2>  
F:351-684/Domain: ferroxidase repeat homology <FO2>  
F:592-1573/Domain: B <DOB>  
F:1183-1461/Region: 9-residue repeats (Q-X-T/N-L-S-P-D-L-S)  
F:1574-2224/Product: coagulation factor Va light chain #status experimental <VAL>  
F:1574-1905/Domain: A3 <DA3>  
F:1581-1905/Domain: ferroxidase repeat homology <FO3>  
F:1667-1765/Region: phospholipid binding #status predicted  
F:1906-2064/Domain: C1 <DC1>  
F:1906-2064/Domain: discoidin I amino-terminal homology <DNI>  
F:2065-2224/Domain: C2 <DC2>  
F:2065-2221/Domain: discoidin I amino-terminal homology <DN2>  
F:51,55,229,297,460,468,554,741,752,760,776,782,821,938,977,1074,1083,1103,1106,1479,  
F:167-193,248-329,500-526,603-684,1125-1751,1907-2061,2066-2221/Disulfide bonds: #sta  
F:334-335/Cleavage site: Arg-Asn (protein C) #status predicted  
F:363,693,1546/Binding site: sulfate (Tyr) (covalent) #status predicted  
F:376-377/Cleavage site: Arg-Ser (coagulation factor Xa) #status predicted  
F:382,1338/Binding site: carboxylate (Asn) (covalent) #status absent  
F:534-535/Cleavage site: Arg-Gly (protein C) #status predicted  
F:737-738/Cleavage site: Arg-Ser (coagulation factor Xa, thrombin) #status experiment  
F:1046-1047/Cleavage site: Arg-Thr (coagulation factor Xa, thrombin) #status experime  
F:1573-1574/Cleavage site: Arg-Ser (thrombin) #status experimental

Query Match 30.6%; Score 2406.5; DB 1; Length 2224;  
Best Local Similarity 26.3%; Pred. No. 5,4e-155;  
Matches 599; Conservative 279; Mismatches 48; Indels 909; Gaps 37;  
Qy 22 RRYVLGAVELSDMDYQSDGLCELPVDAFPPRPVPSKSPFTSVV-YKKTLEFVYHLENI 80  
Db 32 KQYVVAAGISMSYRPE-----PTNSLMLSTSRKVIYRREYR-FKK 75  
Qy 81 AKDRPMMGLLPTIOAEYDVTVITLKNMASHPVSLAVGVSYKMSGAEYDDQTSOR 140  
Db 76 EKQSTISGLLPTLAEVGDILKVFHKNKADKPLSIHQGIRYSKLSGASYLDFRPA 135  
Qy 141 EKEDKVPFGSGSTTYWQYLKENGPMASDPCLCTLYLEHVDLMDNLGLGALLVCRE 200  
Db 136 EKMDADVAPRGRETYEWSISEDSPTPHDPPCLTTHLYSHENMLDEPNSGLGLPLDCK 195  
Qy 201 GSTLAKKETO-TLHK-FLILFAVPEDEKSWHSETKNSLMQDRDASARAPKMHNTGVYN 258  
Db 196 GTLTGEGTKTPKQYLVLFAVPEDEKSWHSETKNSLMQDRDASARAPKMHNTGVYN 239  
Qy 259 RSLPLGLIGCHRSKVWNVIGMTTPREVHSIFLEGHTFLVRNHRQASLEISPTFLTAOTL 318  
Db 240 GTMPDITVCAHDIHSHLLGMSGRLPFIHNGQVLEQNHKVSATLVASTSTTANMT 299  
Qy 319 LMDLGOFLFLFCHTISHQNDQMEAYKVDSCPREPDRKMKNNDEADYDDDLTDEMDYVR 378  
Db 300 VGEGRKWIISLTPKHIQAGQAYADIKKCPKTRMLKTFRE----- 342  
Qy 379 FDDDNPSFPIQIRSVAKKPKPMVNHIAEEDDMYAPVLPDRRSYKSOVLNPGORI 438  
Db 343 -----QRRHMKREYFLIAAEVIMDYAPVAPNMMKRYRSOHLNFSMOI 387



QY	439	GRKTKKRYEAAUYDEPTEKTEEA	----	ONEGSLGPLYLVEGOTLLIFKKNQASPRUNY	495	
Db	388	GKHKKKMYIYOUEDESP	-TKHTYVNPNNKKEGILLGPIIRAOVRDTLKYTKNNAASRPISTY	446		
QY	496	PHGIT	-----	DVNPRLYSRLPKGVKHLKDFLILPEIKRYKWTYVDEGPTKSDPCLY	549	
Db	447	PHGVTFSPYEDEVNSPT	-----	SGRNNMTYIRAOVPEETYTKYKNLLEPFEPEENAOCLT	502	
QY	550	RYSSPYNMERDLASGLIGLPLLCYKESVDQROGNOIMSKRANVLFSYFEDENRSWYLTEN	609			
Db	503	RPYUSDVIMDRDIASGLIGLILLICKSRSLDRNGIQRAADIEOOAVFAVDEKNSWYLEDN	562			
QY	610	IQRLLPACQOULEDBREFQASNMIMHSTINGVPSUL	-QLSVCLEHAAYWYILISGATOPL	668		
Db	563	INKEFENDEVKRDPKFEYESNIMSTINGVPSITITTLGFCFDDTVQWHFCSVGTONEIL	622			
QY	669	SVFSPGYTFKKMYEDTLLLPFGSGFTVMSMENGLMILCSHNS	-----	714		
Db	623	TIHTGHSFTYKGRKHEDTLLPFMRGSESYVTYDNGVTMLISMSNSPASKRLRKLFKDY	682			
QY	715	-----	DFRNR	-----	719	
Db	683	KCIPDDEDESYEIFEPEPESTVMATRKMHDRLEPEDESDADYQORLAALAGISFRNS	742			
QY	720	-----	GMTAL	-----	724	
Db	743	SLNOEEEFMLTALALENGTEFYVSSMTDIIIVGSNYSPPNISKFTYVNNLAEPOKAPSHQ	802			
QY	725	-----	LKSSCDKNFTGDYUEDESD	-----	744	
Db	803	ATTAGSPRLRLIGKNSVLSNSTAHSPTSIEDPIEDPDPDYTGIRLSLGAERKSOEH	862			
QY	745	-----	-----	-----	744	
Db	863	AKHGRPYERDOAAKHREFSWMLLAKHYGRHLSDOTGSPSKRPMEDLPQOTGSPSRMR	922			
QY	745	-----	-----	ISATLSKNN	-----	755
Db	923	PWKDPSPDLLLKQSNSSKTLVGRMHLASEKSYELLIDOTDEDTAVNMMLISPONASRAW	982			
QY	756	-----	-----	-----	-----	755
Db	983	GESTPLANKKQKOSGNKPRRYVNHKSLQVRQDQSKRLKKSOPLIKTRKKKKEKHTHNP	1042			
QY	756	IEPRSF	-----	-----	-----	763
Db	1043	LSRPTFHRLSEAVNTEFSERKLHSLVLKXSNSTSLPDTLNOITLPSMDGWTIASLRLDNO	1102			
QY	764	NSRH	-----	-----	-----	771
Db	1103	NSSMDTQOASCPRGLXQTVPRRHHYOTFPIODRQWHSTSDPSHRSSPELSEMLEYDRS	1162			
QY	772	QKQ	-----	-----	-----	785
Db	1163	HKSFPDTSOMSPSSSEHEWQIVISPDLSQVTLSPELSQTNLSPDLSTHTTISPELIQRNL	1222			
QY	786	-----	-----	-----	-----	785
Db	1223	SPALGQMPISPDLSHTTTLSPDLSDHTTTLSDLSQTNLSPELSQTNLSPALGQMLPSDLSH	1282			
QY	786	-----	-----	-----	-----	785
Db	1283	TTLSIDFSQTNLSPELSHMTLSPELSQTNLSPALGQMPISPDLSHTTTLSDFSQTNLSPE	1342			
QY	786	-----	-----	-----	-----	801
Db	1343	LSQTNLSPALGQMLPSDPSHTTTLSDLSQTNLSPELSQTNLSPDLSEMPLEFADLSQIRL	1402			
QY	802	-----	-----	-----	-----	801
Db	1403	TPDLQWTLSPDLGFTDLSNPFQMSLSDLSQVTLSPDISDTLLPDLSQLSPRPDDQ	1462			
QY	802	-----	-----	-----	-----	806

Db 1463 IFYPBESSQSLLLQGFENSFYPLDGLGMPSSPFLNDTFLSKEPNPLVYIGLSKDGIDY 1522

QY 807 ISVENKKE-----DEDIYDE-----DENOSP-----RSFOKTRKH 837

Db 1523 IEIIPKEVQSSSEDYAEIDYVYPDDPKYDRTGNNSSRDPDNIAMYLKSNNGNRRNT 1582

QY 838 FIAAVERLMDYGMSSPHVLNKR--AQSASVPQ--FKVYFOEFTDGSFPOLYRGELN 892

Db 1583 YIAAEIEMDY----SEFVQRETDIEDDDIPEDTITKVVYFKRYLDTFFKRPDREGE 1638

QY 893 EHLGLIGYIYAEVEDNIMVFRNQAASRPYSFSLSYE-----EDRQGAEPKRN 944

Db 1639 EHLGLIGYIYAEVEDNIVQVRFKKNLASRYSILNHAAGLSYKSSGKTYEDDSPFMKEDN 1698

QY 945 FVKRPETTYFWKQOHNAHAPTKDEFDCAKAAYSFSDVDKEVNHSGILGRLVYCHTNTLNP 1004

Db 1699 AVQPRSSSTYYWNAHTERGRSPSGSACRAMYAYAVNNEKRIHSGILGRLITQKGLIHK 1758

QY 1005 AHGRVTVQERALEFTTFDETKSWYFTENMRNCRAPCNIOMEDPTREKYNRPAHNGYI 1064

Db 1759 DSNMPLDRREFVLFTMETDEKKSWEYKSSRSNR-----LTSEMKKSHEPAINQMI 1812

QY 1065 MDTLPLGLVMADDQRIRWLLSMGSNENHSHFEGHVTYAKKREYKALYNLYRQVET 1124

Db 1813 Y-SLPGILMYQDEWVRRLHLNLIGSDIHVHFHQQLTLENGKQHOHGLVWPLLPGSFET 1871

QY 1125 VEMLPKAGIRVRECELIENHLAGASTLELVYYSNKCQPLTMAAGHINDPQITASAGYQ 1184

Db 1872 LEMKASKGWWMLNTVEGNORACMQPFLIMDDCRKPKMLSTGIIISDSIKASFEPLY 1931

QY 1185 WAPKLARLHYGGSINAMSTKE---PFS---WIKVDLAPMIHIGIKTOGAKOFSSLYIS 1238

Db 1932 WEPRLARLNNGSGSYNAVEKLAIEFASKPVIQVDMQEVYITIGTOGAKNHLYKCYTT 1991

QY 1239 QGITMYSLDGGKMQQTYRNSNGTLMVFRGAVDSSGKIHNFNPRLIARYLHPTHSIR 1298

Db 1992 EFYVAYSNQJNMQIFKNSSTRNMYENGNSDASTIKENQDPRIYAKYIISPTAYNR 2051

QY 1299 STLMEMLMGCDLNCSCMLPQMESKAISDAQDTASSYFTNMFAT--WSPSKARLHLOGRNA 1357

Db 2052 PTLRLLELQGEVNCSTPILGMEKNGIKENKQITASSFKKSMWGDVWEPFRALNMQGRVNA 2111

QY 1358 WRPOVNNPEKMWLQVDFQTKMVTGVTTOGVKSLTSMYKEFLISSQDGHQWLFQNG 1417

Db 2112 WQAKANNKQWLEIDLKIKKITAITQGCXSLSEMYVKSXTYITHYSEQGVEMKPYRLKS 2171

QY 1418 KV--KVQGNODSTTPPVVNSLDPELRLRYLRIHQSWHQAALAMEVLGCRAOLY 1471

Db 2172 SMVDKIEFGNNTNGHVKNFENPPIISRFIRKVIERTWQSIARLRELGQC---DIY 2224

RESULT 5

T42764

C:Species: Mus musculus (house mouse)

C:Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 09-Jun-2000

C:Accession: T42764

R:Yang, T.L.; Cui, J.; RehmTulla, A.; Yang, A.; Mousaelli, M.; Kaufman, R.J.; Ginsbu

Blood 91, 4593-4599, 1998

A:Title: The structure and function of murine factor V and its inactivation by protei

A:Reference number: 222270; MUID:98282202

A:Accession: T42764

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-2183 <AN>

A:Cross-references: EMBL:U529925; NID:g3219690; PID:g3219691; PIDN:AAC99553.1

C:Function:

A:Pathway: blood coagulation

C:Superfamily: coagulation factor V; discoidin I amino-terminal homology; ferroxidase

C:Keywords: blood coagulation; duplication; glycoprotein; phospholipid binding; plasm

F;350-682/domain: ferroxidase repeat homology <FOX1>

F;1541-1864/domain: ferroxidase repeat homology <FOX2>







[illegible]

F:1564-1565/Cleaveage site: Arg-Ser (thrombin) #status experimental

Query Match 30.2% Score 2379; DB 1; Length 2211;  
 Best Local Similarity 26.3%; Pred. No. 4e-153;  
 Matches 597; Conservative 267; Mismatches 496; Indels 910; Gaps 37;

OY	22	RRVYLGAVELASMDYQMSDLGELPVDAREPPRPKSPFPNTSSVYKTLFVEFTVLEFNIA	81
DB	32	RQFYVAQASIRKNNY-----PESTHL-----SKPPETS--FKKIYREYEA-FOKE	76
OY	82	KRPBPWGLIGFTQIAEYDVTVITLKNMASHPVSLAAVGSYWKASGCAEYDQTSORE	141
DB	77	KPOSTSGILGFTLAEYGDIMKVFHKKKAKPLISHAQIKYKSFSGASVDHTLPE	136
OY	142	KEDDAVFFGGSTTYWQVLKENGPMASPRCLITYSLVSHVDLYKDLNGLIGALLVCBEG	201
DB	137	KMDDAVAGQETTYEWTIIEHSGPPHDDPCLTIITYYVNLVEDFNGSLILCLCKG	196
OY	202	SLAKETQOTL--HKFTLLFAVPEDECKSWHSETKNSLMDRDAASARAPKHHYNGYNNR	259
DB	197	TLTEGTOKMEKQVNLFAVPEDECKSWHSETKNSLMDRDAASARAPKHHYNGYNNR	240
OY	260	SLPGLIGCHRRSVYVHWITGMGTTPREVHSIFLEGHTFLVRNHRQASLEISPTFTFLAOTLL	319
DB	241	TMPTITVCANHDHISWHLIGMSGPELFSIHFGQVLEQNHKISATILVTSATSTANMTV	300
OY	320	MDLGQFLFCHSSHQHDGMEVYVDSCPREPOLRMKNNEBAEDYDDDLTDESDVYRF	379
DB	301	SPEGMFTLASLPRHFQGMQAYIDIKCAKTKRPMK-----LTRDQ-----	343
OY	380	DDDNSPSTQIRSVAKKPKPTVWHYIAEEEDMDVAPLYLARDPSYKSOYLNNQPOIG	439
DB	344	-----RRHKRMETFLAEEDYIMDAPILPAMDKTKRSLHDLNFSNIG	388
OY	440	KRYKVRMAAYTDEFKRR--EALQHEGSLIGPLLYCGVDTLLIFKNQASRPYNTYRPH	497
DB	389	KHYKVVVQYQODSEFTRLRDEPSESGGILGPIIRAQVROTLLKVFENMASRGSITRPH	448
OY	438	GIT-----DVAPLYSRLLPKGVKHLKDPPLPGEIFYKMTVYVEDGPTKSDPCL	548
DB	449	GVTESPVDNEVSSSTGSSNTMIRAVR-----PGETTYRKNNILLESDEPTEDAQCL	500
OY	549	TRYYSFPMNRDLASGLIPLLITCKRKYSDORGNQMSDKRNVILSPENPNSWYTE	608
DB	501	TRPYTSNDITRDLASGLIGLLICKSRSLDRGIRADIEQDAVFAPFENKSWIYED	560
OY	609	NIQRFLLPDPACVQLEDEPFOASNIHM-----SINGVYDSIQ--LSVCIEAAYVYTLISG	662
DB	561	NIYKCEMPKRYKRDPRKFYESNIMSNTLTAINGYVESILPILGFCDDOIVQNHFCVSG	620
OY	663	AQDFLVSFSGYTFKHKMYEDTLTFPESGETVEMKSNPGLMTIGCHNSDFRNKMT	722
DB	621	TQNDLITLHFHGHSTFYKRRHEDTLTPRMQGESYTVMDVGTWMLTTMNSNRSRKLIR	680
OY	723	ALLKVSQCDKMTGD-----YYEDS-----YEDISAYL---	749
DB	661	LRFRAKTRINDDDSEYIIEPSGSTAMTKKLIHDSSEIENDADSDYODELALLIGL	740
OY	750	-----	749
DB	741	RSFRNSSLNQKDELNTALALEKDSERTIPPSANRSLSDNSSSSRHSRLAKKPAESLK	800
OY	750	-----LSKNNAIEP-----	758
DB	801	TLHLDEAPASPLEHAGLDKNSALNPPMAHSSPYSEDREDHPLSDVTGVSLLPGTG	860
OY	759	-----RSFSQ-----	763
DB	861	EKNRPAKHQRFQYRGQAAKHKFSQTRFPAKHKTTRLSDQNSSSSRMGPMEDIPSDLL	920
OY	764	-----	763



Dd	921	LQCKDPYKILNGEHNHLYSEKSGYELIQDANENKTYNKLPSNRPQDSRFGENIPFKNSHG	980
Qy	764	-NSRHPs-----TROKOFNAP-----PVL	782
Dd	981	KOSGHPFLVTRRRKPLRODRDRNSRLKEGLPLRIKTRRKKKEEKPAHVPLSPRSFNBRL	1040
Qy	783	-----KRHOEi-----TRTLOSDOEi-----DYDD-----	805
Dd	1041	GEVNASFSDRRHNSHLHASNETSIDLNOTPPSMLSLAASLPBDHOTSPNDTTSOT	1100
Qy	806	-----TISVEKKEDFDIYDED-----ENOSP-----	827
Dd	1101	SSPDLPLPYVSPEBHGYIFPFIQDSDDPHHSTTAPSNRSPDPTHSTTAPSNRSPPTPSQPL	1160
Qy	828	-----	827
Dd	1161	NYDLNRNRAIPTDVSQLFPSELEWQATATSLDSQPSISPDLGOMALSPDGGQESLSPDL	1220
Qy	828	-----	827
Dd	1221	GQTSPLDLSQESLSPDLGQTALSPPDSQESLSPDLGQTALSPPDSQESLSPDLGQTALS	1280
Qy	828	-----	827
Dd	1281	PDRGQESLSPDLQOTSLSPDLSSQESLSPDLGQTALSPPDSQESLSPDLGQTALSPPDQ	1340
Qy	828	-----	827
Dd	1341	SLSPDLGQOTSLSPDLQESLSPDLGQTALSPDPSQESLSPDLGQOTSLSPDLGQESLSPDL	1400
Qy	828	-----	827
Dd	1401	GQTALSPDLSSQESLSPDLGQOTPLSPDLSSLSLSPDLSQLDKQTSPPDLNOTSHTSSESS	1460
Qy	828	-----	827
Dd	1461	QSLPLRPFQGTFFPNAIDGQMPSPRPDSTLNNFTIPEERNPLVYVGLSDDDGDIIEIIPRQ	1520
Qy	828	-----RSFOKTRRHYFLTAVER	844
Dd	1521	KEESSEEDYGEFEFVAANDPYQOTDLRPDINSRNPDIAMAYLRSNNTGNRRKYUYTAABEi	1580
Qy	845	LMDYGNSSSPHVLNRNQSOS---VPO---FKKYVPOEFTGSGTQPLYRBEIENHIGLL	898
Dd	1581	SMWYS-----KEYQSDVDVDPEDDYKKVAFKRYKJDTSTTKLDPQGEYESHGIL	1631
Qy	899	GPIRAEVEDINIVTFERNQASRPYSFYSLSLSE-----EDOROGAEBRKNEVFKNE	950
Dd	1632	GPYIRAEVDVDIQYRFKNLASRPYSLAHGSLSEKSSGKTYEDDSPRMEKEDNALIQNK	1691
Qy	951	TKTYEMKVOHNAAPTKDEFCKAMAYFSDVLEKDVHSGILGIPLLVCHTNTLNPARGHOV	1010
Dd	1692	TYTYVNHATTRSQEPENGSACRAWAYSAVPEKDISHGLIGPLILICRKGTLDKETNMPV	1751
Qy	1011	TYOEFALFTIPEDITSWYTFENNERCGRACNIOEMEDPTFKENYRFAIINGYIMDTLPB	1070
Dd	1752	DMKEFLLFVFPDEKSSWYDYDKRPTSMWRASS-----EYKNSHEFNAINGMETYN-LPG	1804
Qy	1071	LVAAOORIRWYLLSMQSNENISIHPSGVHFTYRKKKEEKYKALYNLPVGFVEYEMPS	1130
Dd	1805	LRRYEDQEWRLHLNLNIGGSDIHVYHNGOTLLENGTQOHLGWAPLLPGSFKTLEMKAS	1864
Qy	1131	KAGINVECEIGHLHAGMSTFLVYSNKCOTPLGMASGHTRDFOITASGOYGOMAPKLA	1190
Dd	1865	KRGWMLLDTEVGEIQRGMOPTPLIYDRBECKMPRGISTGLADSOIQASFEWGWEPRLA	1924
Qy	1191	RHYSGSINAM-----STK---EPSPWIKVDLAPMIHGIKTQGAROKFSSLYISOFTIMY	1244
Dd	1925	RUNNGSGYNAMIKEKLSFEENPEBWIOYDMQKEVLLGICOTOGAKHYLKPYYTTEFCVAY	1984
Qy	1245	SLDGKKMOPYRGASTGTGLAMFFGVNDSSGIKHNIENFPPIARTARYIRHDPHHSINSTLME	1304
Dd	1985	SLDRKMWRIEKGSTRNVMTEFGGSDASTIKENDIDPVAARYIRISTGSYNKPALBLE	2044

Oy 1305 LMGCGLSCSMPLDGESKAISDAQTASSYFPMNPAI--WSPSKARLHLHGGSNMARPQV 1363  
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 2045 LQCEYVNCSTPILPGHESGKIEKKQITTAASSEFKSSWMGNWTPEPLRLANOGRYNMWQAKAN 2104  
  
Oy 1364 NPKEMVLQDFDKTMKVTVGTTQGVKSLLTSMYKEEPLLSSODSQHWTLFFONGRV--KV 1421  
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 2105 NNNGWLQIDLLTKIKRTAIVTGCKSLSEMSWVKSYTHYSDDGTDMKPYPEKSSMDVKI 2164  
  
RESULT 7  
KUUU  
ferroxidase (EC 1.16.3.1) precursor [validated] - human  
N:Alternate names: ceruloplasmin  
N:Contains: ferroxidase long form (CP-1); ferroxidase short form (CP-2)  
C:Species: Homo sapiens (man)  
C:Date: 31-Aug-1980 #sequence.revision.12-May-1995 #text.change.08-Dec-2000  
C:Accession: A25443; A24165; A35450; A00524; I59067  
R:Koschinsky, M.L.; Funk, W.D.; van Oost, B.A.; Macgillivray, R.T.A.  
Proc. Natl. Acad. Sci. U.S.A. 83, 5086-5090, 1986  
A>Title: Complete cDNA sequence of human preceruloplasmin.  
A:Reference number: A25443; MUID:86259737  
A:Accession: A25443  
A:Molecule type: mRNA  
A:Residues: 1-1060,1065-1069 <KOS>  
A:Cross-references: GB:M13699; NID:g180255; PIDN:AAA51976.1; PID:g180256  
A>Note: this is the short or CP-2 alternatively spliced form  
R:Mercer, J.F.B.; Grimes, A.  
FEBS Lett. 203, 185-190, 1986  
A>Title: Isolation of a human ceruloplasmin cDNA clone that includes the N-terminal 1  
A:Reference number: A24165; MUID:86275241  
A:Accession: A24165  
A:Molecule type: mRNA  
A:Residues: 1-40;549-599;784-829;919-952 <MER>  
R:Yang, F.; Friedrichs, W.E.; Cupples, R.L.; Bonifacio, M.J.; Sanford, J.A.; Horton,  
J. Biol. Chem. 265, 10780-10785, 1990  
A>Title: Human ceruloplasmin. Tissue-specific expression of transcripts produced by a  
A:Reference number: A35450; MUID:90285218  
A:Accession: A35450  
A:Molecule type: DNA  
A:Residues: 1007-1064 <YAN>  
A:Cross-references: GB:J05506  
A>Note: this is the long or CP-1 alternatively spliced form  
R:Takahashi, N.; Ortel, T.L.; Putnam, F.W.  
Proc. Natl. Acad. Sci. U.S.A. 81, 390-394, 1984  
A>Title: Single-chain structure of human ceruloplasmin: the complete amino acid sequen  
A:Reference number: A00524; MUID:84119493  
A:Accession: A00524  
A:Molecule type: protein  
A:Residues: 79-9ly and 449-Gly were also found  
R:Yang, F.; Naylor, S.; Lum, J.B.; Cutshaw, S.; McCombs, J.L.; Neberhaus, K.H.; MCG  
Proc. Natl. Acad. Sci. U.S.A. 83, 3257-3261, 1986  
A>Title: Characterization, mapping, and expression of the human ceruloplasmin gene.  
A:Reference number: I59067; MUID:86205876  
A:Accession: I59067  
A>Status: translated from GB/EMBL/DDBU  
A:Molecule type: mRNA  
A:Residues: 218-1069 <RES>  
A:Cross-references: GB:M13536; NID:g180248; PIDN:AAA51975.1; PID:g180249  
C:Comment: Ferroxidase is a blue, plasma alpha2-glycoprotein binding 6-17 copper ions  
C:Comment: In Wilson's disease the plasma levels of ferroxidase are diminished or und  
C:Comment: The three fragment chains are produced spontaneously during purification a  
C:Genetics:  
A:Gene: GDB:CP  
A:Cross-references: GDB:I19069; OMIM:117700  
A:Map position: 3q23-q25  
A:Introns: 1006/3; 1061/1  
A>Note: the list of introns is incomplete







R:Ryan, T.P.; Grover, T.A.; Aust, S.D.  
 Arch. Biochem. Biophys. 293, 1-8, 1992  
 A:Title: Rat ceruloplasmin: resistance to proteolysis and kinetic comparison with human  
 A:Reference number: S21692; M0ID:92117681  
 A:Accession: S21692  
 A:Molecule type: protein  
 A:Residues: 20-29, 'Q':902-910 <RYA>  
 C:Superfamily: ferroxidase; ferroxidase repeat homology  
 C:Keywords: copper; glycoprotein; oxidoreductase; plasma  
 F:1-19/Domain: signal sequence #status predicted <Sig>  
 F:20-1059/Product: ferroxidase #status predicted <AMT>  
 F:23-356/Domain: ferroxidase repeat homology <FO1>  
 F:372-712/Domain: ferroxidase repeat homology <FO2>  
 F:727-1053/Domain: ferroxidase repeat homology <FO3>

Query Match	21.4%; Score 1681; DB 1; Length 1059;
Best Local Similarity	32.0%; Pred. No. 4.3e-106;
Matches	385; Conservative 188; Mismatches 429; Indels 200; Gaps 23;
QY	5 LSTCFLLLRFCFSATRRYYLGAVELSDYMQ-SDIGEL-PVDARPPRPVKSPFNTS 62
DB	6 LSAFLFL-HSLAMTRKHHYIGITLAVMDYASGSEKELISVDTEQSNFYLRNPNDRIG 64
QY	63 VYKTKLPEFVTHLENIKKRPPRMGGLGPTQAEYDPTVYLILKNASHVSIHAAGV 122
DB	65 RYKAKLISEYDGTFTTKTIDKPAWLGFLGPVIAEVDKSVHKNKASRYTHAHGV 124
QY	123 SYWKASEGAEYDQTSQREKEDDKVPFGSGHTYVQWLKENGPMASDPLCLTYLSHYD 182
DB	125 TYTKANEGALYDNTDTPQADDDKLPFGQOYLVLRA-NEPSPGEGDNCYTRISHSD 183
QY	183 LVKDINSGLGALLVCREGSLAKEKTQTL-HKFTLLFAVEDEGKSWHSETKNSLM----- 236
DB	184 APKDIASGLIGLILCKKSLHKEKENIDQEFYLFMSVVDENISWYLEDNIKTCSEPE 243
QY	237 -QDRPAASRRAPKMTGVGVYVRSRLPGLLGCHKSKYVNHVIGTTPREVSIFLEGHTF 295
DB	244 KYDKNEDFQESNRMSINGYTFGSLPGLSMCAEDRKWKVLEFGMENEVDHSALEFGQAL 303
QY	296 LVNRHROASLESPITFLAQTLLMDLGOFLPEHISHQHDGEAVYKVDSCPEEPLR 355
DB	304 TSKNHTDINLEPALLIDVSWAQNPGVWMLSCQNLNLKAGLCAQAFQVDRDC----- 356
QY	356 MKNNEAEYDQDLTDEMDVYRFDSDNSPSFIQIRSAKKHPKTHVYIAAEEEDWDYA 415
DB	357 ---NKPSR-DDDIDQDRHY-----RH-----YIAAEEITWDYA 385
QY	416 P-----LVLPDDRSTYSQYLINNGPQRIGRKKYKKRPMAYTPETF---KTRRL 461
DB	386 PSGDTFTGENTLTSGSDSRVFEQ---GATRGGSKYKLVREYTDTSFNRKRGPD 441
QY	462 QHESGILPLRLEGVDTLLIFKNOASRPYNIYPHGITDVR-----PLYSRLPRGVKL 517
DB	442 EHLGILGRVIAEVDIIRYTHNKGQPLSLDPMGVRFKKEGTYG---PGRSRK 498
QY	518 KDFPLPEIFRYKWTIVTEDEGPTKSDPRLCTIRYSSFVNMRDLASGLIPLLCYES 577
DB	499 QASHVAPKETFYEMTVPRKEMGPTVADPCLSKMYSGVDLTDFGLIGMKICKKS 558
QY	578 VDQRNQIMSKRNVILFSVFDENRSMYLTENIORFLPNPAGVQLEDEPQASNIHSHN 637
DB	559 LLADDRQKDVDEFLYFAVYFENESILLDDNIIMFTAPENVVKEDDEPQSNMHSNN 618
QY	638 GYVFESLQ-LSYCLHEAVAYWILSIGAQDPLSFVFGGYPRKHKMYEDTLPLPFSGPT 696
DB	619 GFMIGNLPCLNLCGESIYWILFSAGNEADVHGIFSGNTYLSGERDQIANLPHKSLT 678
QY	697 VFMSENPGILMILGCHNSDFRNKGATALLKVSQDKWTGDYEDSYEDISAYLSKNNAI 756
DB	679 LLMPDTGESFLVECLITTHYTGCMKQKYTVNQC-----KQFELVYTL----- 722
QY	757 EPRFSQNSRHSTROKQFNAPPVILKRHQREITRTTLQSDQEEIDYDQTLISVEKKEDF 816

Db	723	-----	722
QY	817	DIYDEDNOSPRSFQKTRHPIAVERLMDYGMSSSHVLRNRAQSSV-----	866
Db	723	-----	767
QY	867	---POFKKVVFOEFTDGSFTOPLRGELNENHGLGPIRAVEEDNIWTFRRNOSAPRS	923
Db	768	FIGSKTKKKVYVEEFDSTFREOVKKRAEEHGIIGLPIHADYGVKKVYVKKMATRYS	827
QY	924	FYSLSIYSEEDROGAEPKRNFK--PNEKTYPMKVOHHMPTKDEFDCKAMAYESDV	980
Db	828	IHA-----HGVTKTSSTVAPTLPGEVKTYIMQIPERSGAQTEEDSPICIMAYSTV	877
QY	981	DLEKDVHSLGIPRLVLC---HTNTLNPAGKQVYVQEEALFTTFDEKSKSYFIENNERN	1037
Db	878	DNVKCLISGLIPLLVCRKRSYKVENPK---KKMEFSLFLFVDENESWYLDNDINTY	932
QY	1038	CRAPCNIGMEDPTFKENYRFHAINCYIMDTLPGVLAMADORLRIWLSMGSENENHSIHF	1097
Db	933	PDHPEKVNKNDEEFLIESKNKMAHINKMGKGNLOGLLFMHGVDEVNMYVAMAGNEIDLHYHF	992
QY	1098	SGHVEFTVRKKEEYKMAALNLYPGVEFVEMLPKSKAGIWRVECLIGEHLHAQMSTLFLVYS	1157
Db	993	HGHSTFOYKHRGISHSDVEFEGPTGYQTELEMPQPGTWLHCHVTHIHAGMVTYTVLP	1052
QY	1158	NK 1159	
Db	1053	NQ 1054	
RESULT 9			
A44258			
factor VIII-associated gene B hypothetical protein - human			
C:Species: Homo sapiens (man)			
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 18-Jun-1999			
C:Accession: A44258			
R:Levinson, B.; Kenwright, S.; Gamel, P.; Fisher, K.; Gitschler, J.			
Genomics 14, 585-589, 1992			
A:Title: Evidence for a third transcript from the human factor VIII gene.			
A:Reference number: A44258; MUID:93052386			
A:Accession: A44258			
A:Status: preliminary			
A:Molecule type: mRNA			
A:Residues: 1-216 <LEV>			
A:Cross-references: GB:M0707; NID:g182316; PID:AAA58466.1; PID:g182317			
C:Superfamily: coagulation factor VIII; discoidin I amino-terminal homology; ferroxid			
F:1-53/Domain: discoidin I amino-terminal homology #status atypical <DN1>			
F:57-210/Domain: discoidin I amino-terminal homology <DN2>			
Query Match	14.0%;	Score 1104;	DB 2;
Best Local Similarity	100.0%;	Pred. No. 6.6e-68;	Length 216;
Matches 208;	Conservative 0;	Mismatches 0;	Indels 0;
			Gaps 0;
QY 1264	VEFGVNDSSGIRKHNIFNPPIIARYIRLPHTHYSIRSLRMELMGCDLNSCMLPGMESKA	1323	
Db 9	VEFGVNDSSGIRKHNIFNPPIIARYIRLPHTHYSIRSLRMELMGCDLNSCMLPGMESKA	68	
QY 1324	ISDAQITASSYFTNNFATWSPSKARLHLOGRSNAMPVNNPKEMLOYDFQTKMKVTGYT	1383	
Db 69	ISDAQITASSYFTNNFATWSPSKARLHLOGRSNAMPVNNPKEMLOYDFQTKMKVTGYT	128	
QY 1384	TQGVSLTSMWVKREFLISSSODGHQWTLFQNGKVVKYKFGCONDSFPVNVSLDPLLTR	1443	
Db 129	TQGVSLTSMWVKREFLISSSODGHQWTLFQNGKVVKYKFGCONDSFPVNVSLDPLLTR	188	
QY 1444	YLRHPQSWHVOIALRMEVLGCEADQLY 1471		
Db 189	YLRHPQSWHVOIALRMEVLGCEADQLY 216		



RESULT 10  
 JC4915  
 ags protein precursor - rat  
 N:Alternate names: O-acetyl-cd3 ganglioside.  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 26-Sep-1996 #sequence\_revision 01-Nov-1996 #text\_change 20-Jun-2000  
 C:Accession: JC4915  
 R:Ogura, K.; Nara, K.; Watanabe, Y.; Kohno, K.; Tai, T.; Sanai, Y.  
 Biochem. Biophys. Res. Commun. 225, 932-938, 1996  
 A:Title: Cloning and expression of cDNA for O-acetylation of GD3 ganglioside.  
 A:Reference number: JC4915; MUID:96374422  
 A:Accession: JC4915  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-427 <OGU>  
 A:Cross-references: DDBJ:D84068; NID:q1620006; PIDB:BA12210.1; PID:q1620007  
 A:Experimental source: C57 cell  
 C:Comment: This protein is required for the O-acetylation of disialoganglioside sialic acid  
 C:Genetics:  
 A:Gene: ags  
 C:Superfamily: milk fat globule protein; discoidin I amino-terminal homology; EGF homology  
 F:1-21/Domain: signal sequence #status predicted <SIG>  
 F:28-60/Domain: EGF homology <EG1>  
 F:68-107/Domain: EGF homology <EG2>  
 F:110-267/Domain: discoidin I amino-terminal homology <DN1>  
 F:271-427/Domain: discoidin I amino-terminal homology <DN2>

```
Query Match      8.4%; Score 663; DB 2; Length 427;
Best Local Similarity 40.3%; Pred. No. 2.2e-37;
Matches 142; Conservative 61; Mismatches 123; Indels 26; Gaps 6;
```

OY	1138	ECLIGEHNLHAGMSTFLVYSNK-----CQPLGMAHGHIIDFOITASGOV--	1182
Dd	78	KCLVIEDPDRG--DLFEETICQCPVYGSGIHCELCSTRKLEGGALIDSQSISASYMG	135
OY	1183	---GOMAPKLARLHSSGINAW--STKEPFSEWIKVDLLAPMIHGIRTOGAROKFESSLY	1236
Dd	136	FMLGLRMPGPBLARLRRTGIYNAWMTASSYDSKRPIQVDEFLRKMRVSGVMTQGASRGRAYEY	195
OY	1237	ISOFLIIMSLOGKKMQYTGRGNSTGLTMFFGVNDSSGIAKHINFNPTIARTIRLPHITYS	1296
Dd	196	LKTFVVAVSLSGRRREFIQDESGTDKKEFMQNODNSLKIMNFPLETAAOYIRLPVSCH	255
OY	1297	IRSTRLMEIMGCDLNSCSMPLGMEKSAISDAOTITASSYFT--NMFA-TWSPSKARHLHOG	1353
Dd	256	RGCILRFELLECCELHCSEPRLGCKNNTIPDQOITRNASSSYKTWNLARAFWYTHLGRINDOG	315
OY	1354	RSNAWRPOVNNPKEMLVQDFOKTYMKVTGTTQGYKSLLTSMYVEKLFISSSDGHQHTLF	1413
Dd	316	KINAMTAOSNASKAEWLQYADLTQOKKVGTGITQGFARDFGHIIQVAYASYKAHSDDGQVMTIVY	375
OY	1414	FONGKVKVPFGQODSFTRPVNSLDPRPLLTRLIRLIHPQSWAQIALRMVYLGC	1465
Dd	376	EEQGSUKTFPGQLDNNSHKKNIFEKPEMARVRYRLPSWNRRIITRELLCG	427

C:Keywords: membrane protein  
F:28-60/Domain: EGF homology <EG1>  
F:68-107/Domain: EGF homology <EG2>  
F:140-303/Domain: discoidin I amino-terminal homology <DN1>  
F:307-463/Domain: discoidin I amino-terminal homology <DN2>

[illegible][illegible]



Oy	1297	IRSTLMEIWMGCDJNSNSMPLGMEKSAISDAQITASSYNTN--	-MEATNPSKARJLHOG	1353
Dd	236	RGCTLFFELLGCEISGAEPLGDNTIPNKOITASSFFRTWGLSAFMSWPFYARLDNOG		295
Oy	1354	RSNAMPPOVNNPKEMLVDFQKTKMKVITGVTOGVKSLTSMYVKEPLISSDQGHOWTLF		1413
Dd	296	KENAMTAQNSASASEWLDIDLGSSQRVVTGIITQGAADFCHIQTQYAAAKVAASDDGVSNTEY		355
Oy	1414	PONGKV--KVEQGNODSFTPVNNSLDPLFRTYLRIPHSWQHOLALREVLGC		1465
Dd	356	RQDGALEGKTFPGNLDNNSHKKNMFETPFLTRFVILPVAMHNRILTRDELGC		409

RESULT 13

S65138

N:Glycoprotein antigen MGp57/53, mammary gland - bovine (fragment)

C:Species: Bos primigenius taurus (cattle)

C:Date: 28-Oct-1996 #sequence\_revision 13-Mar-1997 #text\_change 07-Aug-1998

C:Accession: S65138; G48394

R:Aoki, N.; Kishi, M.; Taniguchi, Y.; Adachi, T.; Nakamura, R.; Matsuda, T.

R:Biochim. Biophys. Acta 1245, 385-391, 1995

A:Title: Molecular cloning of glycoprotein antigens MGp57/53 recognized by monoclonal antibodies

A:Reference number: S65138; MUID:56125736

A:Accession: S65138

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: mRNA

R:Residues: 1-401 <AOK>

R:Mathier, I.H.; Banghart, L.R.; Lane, W.S.

R:Biochem. Mol. Biol. Int. 29, 545-554, 1993

A:Title: The major fat globule membrane proteins, bovine components 15/16 and guinea-pig 11-like sequences.

A:Reference number: A48394; MUID:93250576

A:Accession: G48394

A:Status: preliminary

A:Molecule type: protein

R:Residues: 207-220 <MAT>

A:Experimental source: milk

A:Note: sequence extracted from NCBI backbone (NCBI:131457)

C:Superfamily: milk fat globule protein; discoidin I amino-terminal homology: EGF homology

C:Keywords: glycoprotein

F:1-32/Domain: EGF homology (fragment) <EG1>

F:40-79/Domain: EGF homology <EG2>

F:82-239/Domain: discoidin I amino-terminal homology <DN1>

F:243-401/Domain: discoidin I amino-terminal homology <DN2>

[illegible]

Db 346 EYKDPGASESKTFPGMNDNNSHKKNITETPFQARFVRIDPVAWHNRITLRVELLGC 401

RESULT 14

PAS-6/7 protein precursor - bovine  
 C:Species: Bos primigenius taurus (cattle)  
 C:Date: 04-Dec-1997 #sequence-revision 12-Dec-1997 #text-change 20-Jun-2000  
 C:Accession: S74211; S78114; S24181  
 R:Harrigaard, J.; Andersen, M.H.; Berglund, L.; Rasmussen, J.T.; Petersen, T.E.  
 Eur. J. Biochem. 240, 628-636, 1996  
 A:Title: Characterization of glycoprotein PAS-6/7 from membranes of bovine milk fat g  
 A:Reference number: S74211; MUID:97008954

A:Residues: 1-227 <HYA>  
A:Cross-references: EMBL:X91895; NID:g16322778; PIDN:CAA62997.1; PID:g1632779  
A:Accession: S78114  
A:Molecule type: protein  
A:Residues: 19-85;96-110;140-165;174-216;221-232;248-277;285-293;309-337;339-420;425-  
R:Kim, D.H.; Kanno, C.; Mizokami, Y.  
Biochim. Biophys. Acta 1122, 203-211, 1992  
A:Title: Purification and characterization of major glycoproteins, PAS-6 and PAS-7, f  
A:Reference number: S23926; MUID:9235107  
A:Accession: S24181  
A:Molecule type: protein  
A:Residues: 383-394 <KIM>  
C:Superfamily: milk fat globule protein; discoidin I amino-terminal homology; EGF hom  
C:Keywords: blocked amino end; status predicted <SIG>  
F:1-18/Domain: signal sequence #status predicted <SIG>  
F:19-427/Product: PAS-6/7 protein #status experimental <MAT>  
F:24-58/Domain: EGF homology <EG1>  
F:66-105/Domain: EGF homology <EG2>  
F:108-265/Domain: discoidin I amino-terminal homology <DN1>  
F:269-427/Domain: discoidin I amino-terminal homology <DN2>  
F:24-35;29-47;49-58;66-77;71-94;96-105/Disulfide bonds: #status predicted  
F:27/Binding site: carbohydrate (Ser) (covalent) #status experimental  
F:34/Binding site: carbohydrate (Thr) (covalent) #status experimental  
F:59;227/Binding site: carbohydrate (Asn) (covalent) #status experimental  
F:109-265;252-256;270-427/Disulfide bonds: #status experimental

Query Match 8.1k; Score 635; DB 2; Length 427;  
Best Local Similarity 37.9k; Pred. No. 1.8e-35;  
Matches 135; Conservative 66; Mismatches 123; Indels 32; Gaps 7;

[illegible]



A25945  
coagulation factor VIII - pig (fragment)  
C:Species: Sus scrofa domestica (domestic pig)  
C:Date: 16-Aug-1988 #sequence\_revision 16-Aug-1988 #text\_change 21-Jul-2000  
C:Accession: A25945  
R:Toole, J.J.; Pittman, D.D.; Orr, E.C.; Murtha, P.; Wasley, L.C.; Kaufman, R.J.  
Proc. Natl. Acad. Sci. U.S.A. 83, 5939-5942, 1986  
A:Title: A large region (approx 195 kDa) of human factor VIII is dispensable for in vitro  
A:Reference number: A25945; MUID:86287369  
A:Accession: A25945  
A:Status: nucleic acid sequence not shown: not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 1-869 <TOO>  
C:Superfamily: coagulation factor VIII; discoidin I amino-terminal homology; ferroxidase

Query Match 6.98; Score 543; DB 2; Length 869;  
Best Local Similarity 75.2%; Pred. No. 1.1e-28;  
Matches 100; Conservative 14; Mismatches 19; Indels 0; Gaps 0;  
QY 779 PPVLRKHOREITRTTLOSDEIDYDDTISVEMKKEDFDIYDEDNOSPRSPQKTRHYF 838  
DB 737 PPVLRKHORDISLPTROPEDMDYDDIFSTETKGEDFDIGEDNDPDRSFQKTRHYF 796  
QY 839 IAAVERLMDYGMSSSPHYLRNRAQSGSVPOFKVVFQETDGSFTQPLYRGELNEHLGL 898  
DB 797 IAAVEQLMDYGMSESPRALRNRAQNGEVPRFKVYFRERADGSFTNPSYRGELNKHGL 856  
QY 899 GPYIRAEVEDNIM 911  
DB 857 GPYIRAEVEDNIM 869

Search completed: November 17, 2001, 13:23:23  
Job time: 72 sec



**THIS PAGE BLANK (USPTO)**